

SEQUENCE LISTING

- (1) GENERAL INFORMATION:
- (i) APPLICANT: Iris Pecker, Israel Vlodavsky and Elena Feinstein
 - (ii) TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY AND EXPRESSION OF SAME IN GENETICALLY MODIFIED CELLS
 - (iii) NUMBER OF SEQUENCES: 47
 - (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
 - (B) STREET: 2001 Jefferson Davis Highway, Suite 207
 - (C) CITY: Arlington
 - (D) STATE: Virginia
 - (E) COUNTRY: United States of America
 - (F) ZIP: 22202
 - (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
 - (B) COMPUTER: Twinhead* Slimnote-890TX
 - (C) OPERATING SYSTEM: MS DOS version 6.2, Windows version 3.11
 - (D) SOFTWARE: Word for Windows version 2.0 converted to an ASCII file
 - (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
 - (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/922,170
 - (B) FILING DATE: 2 SEP 1997
 - (A) APPLICATION NUMBER: 09/109,386
 - (B) FILING DATE: 10 JUL 1998
 - (A) APPLICATION NUMBER: PCT/US98/17954
 - (B) FILING DATE: 31 AUG 1998
 - (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Friedmam, Mark M.
 - (B) REGISTRATION NUMBER: 33,883
 - (C) REFERENCE/DOCKET NUMBER: 910/14
 - (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 972-3-5625553
 - (B) TELEFAX: 972-3-5625554
 - (C) TELEX:
- (2) INFORMATION FOR SEQ ID NO:1:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
CCATCCTAAT ACGACTCACT ATAGGGC 27
- (2) INFORMATION FOR SEQ ID NO:2:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GTAGTGATGC CATGTAAGTGAATC 24

- 5

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1721
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTAGAGCTTT CGACTCTCCG CTGCGCGGCA GCTGGCGGGG GGAGCAGCCA GGTGAGCCCA 60
 AGATGCTGCT GCGCTCGAAG CCTGCGCTGC CGCCGCCGCT GATGCTGCTG CTCCTGGGGC 120
 CGCTGGGTCC CCTCTCCCTT GGCGCCCTGC CCCGACCTGC GCAAGCACAG GACGTCGTGG 180
 ACCTGGACTT cTTCACCCAG GAGCGCTGC ACCTGGTGAG CCCCTCGTTC CTGTCCGTCA 240
 CCATTGACGC CAACCTGGCC ACGGACCCGC GGTTCCTCAT CCTCTGGGT TCTCCAAAGC 300
 TTCGTACCTT GGCCAGAGGC TTGTCTCCTG CGTACCTGAG GTTGGTGGC ACCAAGACAG 360
 ACTTCCTAAT TTTGATCCC AAGAAGGAAT CAACCTTTGA AGAGAGAAGT TACTGGCAAT 420
 CTCAAGTCAA CCAGGATATT TGCAATATG GATCCATCCC TCCTGATGTG GAGGAGAAGT 480
 TACGGTTGGA ATGGCCCTAC CAGGAGCAAT TGCTACTCCG AGAACACTAC CAGAAAAAGT 540
 TCAAGAACAG CACCTACTCA AGAAGCTCTG TAGATGTGCT ATACACTTTT GCAAACTGCT 600
 CAGGACTGGA CTTGATCTTT GGCCTAAATG CGTTATTAAG AACAGCAGAT TTGCAGTGGA 660
 ACAGTTCTAA TGCTCAGTTG CTCCTGGACT ACTGCTCTTC CAAGGGGTAT AACATTTCTT 720
 GGGAACTAGG CAATGAACCT AACAGTTTCC TTAAGAAGGC TGATATTTTC ATCAATGGGT 780
 CGCAGTTAGG AGAAGATTAT ATTCAATTGC ATAACTTCT AAGAAAGTCC ACCTTCAAAA 840
 ATGCAAACT CTATGGTCTT GATGTTGGTC AGCCTCGAAG AAAGACGGCT AAGATGCTGA 900
 AGAGCTTCCT GAAGGCTGGT GGAGAAGTGA TTGATTCACT TACATGGCAT CACTACTATT 960
 TGAATGGACG GACTGCTACC AGGGAAGATT TTCTAAACCC TGATGTATTG GACATTTTTA 1020
 TTTTATCTGT GCAAAAAGTT TTCCAGGTGG TTGAGAGCAC CAGGCCTGGC AAGAAGGTCT 1080
 GGTAGGAGA AACAAGCTCT GCATATGGAG GCGGAGCGCC CTTGCTATCC GACACCTTTG 1140
 CAGCTGGCTT TATGTGGCTG GATAAATTGG GCCTGTCAGC CCGAATGGGA ATAGAAGTGG 1200
 TGATGAGGCA AGTATTCTTT GGAGCAGGAA ACTACCATTT AGTGGATGAA AACTTCGATC 1260
 CTTTACCTGA TTATTGGCTA TCTCTTCTGT TCAAGAAATT GGTGGGCACC AAGGTGTTAA 1320
 TGGCAAGCGT GCAAGGTTCA AAGAGAAGGA AGCTTCGAGT ATACCTTCAT TGCACAAACA 1380
 CTGACAATCC AAGGTATAAA GAAGGAGATT TAACTCTGTA TGCCATAAAC CTCCATAACG 1440
 TCACCAAGTA CTTGCGGTTA CCCTATCCTT TTTCTAACAA GCAAGTGGAT AAATACCTTC 1500
 TAAGACCTTT GGGACCTCAT GGATTACTTT CCAAATCTGT CCAACTCAAT GGTCTAACTC 1560
 TAAAGATGGT GGATGATCAA ACCTTGCCAC CTTTAATGGA AAAACCTCTC CGGCCAGGAA 1620
 GTTCACTGGG CTTGCCAGCT TTCTCATATA GTTTTTTGT GATAAGAAAT GCCAAAGTTG 1680
 CTGCTTGCACT CTGAAAATAA AATATACTAG TCCTGACACT G 1721

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 543
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Leu Leu Arg Ser Lys Pro Ala Leu Pro Pro Pro Leu Met Leu Leu
 5 10 15

Leu Leu Gly Pro Leu Gly Pro Leu Ser Pro Gly Ala Leu Pro Arg Pro
 20 25 30

Ala Gln Ala Gln Asp Val Val Asp Leu Asp Phe Phe Thr Gln Glu Pro
 35 40 45

Leu His Leu Val Ser Pro Ser Phe Leu Ser Val Thr Ile Asp Ala Asn
 50 55 60

Leu Ala Thr Asp Pro Arg Phe Leu Ile Leu Leu Gly Ser Pro Lys Leu

65	70	75	80
Arg Thr Leu Ala Arg Gly Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly			
85	90	95	
Thr Lys Thr Asp Phe Leu Ile Phe Asp Pro Lys Lys Glu Ser Thr Phe			
100	105	110	
Glu Glu Arg Ser Tyr Trp Gln Ser Gln Val Asn Gln Asp Ile Cys Lys			
115	120	125	
Tyr Gly Ser Ile Pro Pro Asp Val Glu Glu Lys Leu Arg Leu Glu Trp			
130	135	140	
Pro Tyr Gln Glu Gln Leu Leu Leu Arg Glu His Tyr Gln Lys Lys Phe			
145	150	155	160
Lys Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Val Leu Tyr Thr Phe			
165	170	175	
Ala Asn Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu Asn Ala Leu Leu			
180	185	190	
Arg Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu Leu Leu			
195	200	205	
Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile Ser Trp Glu Leu Gly Asn			
210	215	220	
Glu Pro Asn Ser Phe Leu Lys Lys Ala Asp Ile Phe Ile Asn Gly Ser			
225	230	235	240
Gln Leu Gly Glu Asp Tyr Ile Gln Leu His Lys Leu Leu Arg Lys Ser			
245	250	255	
Thr Phe Lys Asn Ala Lys Leu Tyr Gly Pro Asp Val Gly Gln Pro Arg			
260	265	270	
Arg Lys Thr Ala Lys Met Leu Lys Ser Phe Leu Lys Ala Gly Gly Glu			
275	280	285	
Val Ile Asp Ser Val Thr Trp His His Tyr Tyr Leu Asn Gly Arg Thr			
290	295	300	
Ala Thr Arg Glu Asp Phe Leu Asn Pro Asp Val Leu Asp Ile Phe Ile			
305	310	315	320
Ser Ser Val Gln Lys Val Phe Gln Val Val Glu Ser Thr Arg Pro Gly			
325	330	335	
Lys Lys Val Trp Leu Gly Glu Thr Ser Ser Ala Tyr Gly Gly Gly Ala			
340	345	350	
Pro Leu Leu Ser Asp Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys			
355	360	365	
Leu Gly Leu Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val			

098813 1101

370	375	380	
Phe Phe Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro			
385	390	395	400
Leu Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr			
	405	410	415
Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu Arg			
	420	425	430
Val Tyr Leu His Cys Thr Asn Thr Asp Asn Pro Arg Tyr Lys Glu Gly			
	435	440	445
Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr Lys Tyr Leu			
	450	455	460
Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp Lys Tyr Leu Leu			
	465	470	475
Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys Ser Val Gln Leu Asn			
	485	490	495
Gly Leu Thr Leu Lys Met Val Asp Asp Gln Thr Leu Pro Pro Leu Met			
	500	505	510
Glu Lys Pro Leu Arg Pro Gly Ser Ser Leu Gly Leu Pro Ala Phe Ser			
	515	520	525
Tyr Ser Phe Phe Val Ile Arg Asn Ala Lys Val Ala Ala Cys Ile			
	530	535	540
			543
(2)	INFORMATION FOR SEQ ID NO:11:		
(1)	SEQUENCE CHARACTERISTICS:		
(A)	LENGTH:	1721	
(B)	TYPE:	nucleic acid	
(C)	STRANDEDNESS:	double	
(D)	TOPOLOGY:	linear	
(xi)	SEQUENCE DESCRIPTION: SEQ ID NO:11:		
		CT AGA GCT TTC GAC	14
TCT CCG CTG CGC GGC AGC TGG CGG GGG GAG CAG CCA GGT GAG CCC AAG			62
ATG CTG CTG CGC TCG AAG CCT GCG CTG CCG CCG CCG CTG ATG CTG CTG			110
Met Leu Leu Arg Ser Lys Pro Ala Leu Pro Pro Pro Leu Met Leu Leu	5	10	15
CTC CTG GGG CCG CTG GGT CCC CTC TCC CCT GGC GCC CTG CCC CGA CCT			158
Leu Leu Gly Pro Leu Gly Pro Leu Ser Pro Gly Ala Leu Pro Arg Pro	20	25	30
GCG CAA GCA CAG GAC GTC GTG GAC CTG GAC TTC TTC ACC CAG GAG CCG			206
Ala Gln Ala Gln Asp Val Val Asp Leu Asp Phe Phe Thr Gln Glu Pro	35	40	45
CTG CAC CTG GTG AGC CCC TCG TTC CTG TCC GTC ACC ATT GAC GCC AAC			254

Leu His Leu Val Ser Pro Ser Phe Leu Ser Val Thr Ile Asp Ala Asn
 50 55 60

CTG GCC ACG GAC CCG CGG TTC CTC ATC CTC CTG GGT TCT CCA AAG CTT 302
 Leu Ala Thr Asp Pro Arg Phe Leu Ile Leu Leu Gly Ser Pro Lys Leu
 65 70 75 80

CGT ACC TTG GCC AGA GGC TTG TCT CCT GCG TAC CTG AGG TTT GGT GGC 350
 Arg Thr Leu Ala Arg Gly Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly
 85 90 95

ACC AAG ACA GAC TTC CTA ATT TTC GAT CCC AAG AAG GAA TCA ACC TTT 398
 Thr Lys Thr Asp Phe Leu Ile Phe Asp Pro Lys Lys Glu Ser Thr Phe
 100 105 110

GAA GAG AGA AGT TAC TGG CAA TCT CAA GTC AAC CAG GAT ATT TGC AAA 446
 Glu Glu Arg Ser Tyr Trp Gln Ser Gln Val Asn Gln Asp Ile Cys Lys
 115 120 125

TAT GGA TCC ATC CCT CCT GAT GTG GAG GAG AAG TTA CGG TTG GAA TGG 494
 Tyr Gly Ser Ile Pro Pro Asp Val Glu Glu Lys Leu Arg Leu Glu Trp
 130 135 140

CCC TAC CAG GAG CAA TTG CTA CTC CGA GAA CAC TAC CAG AAA AAG TTC 542
 Pro Tyr Gln Glu Gln Leu Leu Leu Arg Glu His Tyr Gln Lys Lys Phe
 145 150 155 160

AAG AAC AGC ACC TAC TCA AGA AGC TCT GTA GAT GTG CTA TAC ACT TTT 590
 Lys Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Val Leu Tyr Thr Phe
 165 170 175

GCA AAC TGC TCA GGA CTG GAC TTG ATC TTT GGC CTA AAT GCG TTA TTA 638
 Ala Asn Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu Asn Ala Leu Leu
 180 185 190

AGA ACA GCA GAT TTG CAG TGG AAC AGT TCT AAT GCT CAG TTG CTC CTG 686
 Arg Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu Leu Leu
 195 200 205

GAC TAC TGC TCT TCC AAG GGG TAT AAC ATT TCT TGG GAA CTA GGC AAT 734
 Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile Ser Trp Glu Leu Gly Asn
 210 215 220

GAA CCT AAC AGT TTC CTT AAG AAG GCT GAT ATT TTC ATC AAT GGG TCG 782
 Glu Pro Asn Ser Phe Leu Lys Lys Ala Asp Ile Phe Ile Asn Gly Ser
 225 230 235 240

CAG TTA GGA GAA GAT TAT ATT CAA TTG CAT AAA CTT CTA AGA AAG TCC 830
 Gln Leu Gly Glu Asp Tyr Ile Gln Leu His Lys Leu Leu Arg Lys Ser
 245 250 255

ACC TTC AAA AAT GCA AAA CTC TAT GGT CCT GAT GTT GGT CAG CCT CGA 878
 Thr Phe Lys Asn Ala Lys Leu Tyr Gly Pro Asp Val Gly Gln Pro Arg
 260 265 270

AGA AAG ACG GCT AAG ATG CTG AAG AGC TTC CTG AAG GCT GGT GGA GAA 926
 Arg Lys Thr Ala Lys Met Leu Lys Ser Phe Leu Lys Ala Gly Gly Glu

275

280

285

GTG ATT GAT TCA GTT ACA TGG CAT CAC TAC TAT TTG AAT GGA CGG ACT 974
 Val Ile Asp Ser Val Thr Trp His His Tyr Tyr Leu Asn Gly Arg Thr
 290 295 300

GCT ACC AGG GAA GAT TTT CTA AAC CCT GAT GTA TTG GAC ATT TTT ATT 1022
 Ala Thr Arg Glu Asp Phe Leu Asn Pro Asp Val Leu Asp Ile Phe Ile
 305 310 315 320

TCA TCT GTG CAA AAA GTT TTC CAG GTG GTT GAG AGC ACC AGG CCT GGC 1070
 Ser Ser Val Gln Lys Val Phe Gln Val Val Glu Ser Thr Arg Pro Gly
 325 330 335

AAG AAG GTC TGG TTA GGA GAA ACA AGC TCT GCA TAT GGA GGC GGA GCG 1118
 Lys Lys Val Trp Leu Gly Glu Thr Ser Ser Ala Tyr Gly Gly Gly Ala
 340 345 350

CCC TTG CTA TCC GAC ACC TTT GCA GCT GGC TTT ATG TGG CTG GAT AAA 1166
 Pro Leu Leu Ser Asp Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys
 355 360 365

TTG GGC CTG TCA GCC CGA ATG GGA ATA GAA GTG GTG ATG AGG CAA GTA 1214
 Leu Gly Leu Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val
 370 375 380

TTC TTT GGA GCA GGA AAC TAC CAT TTA GTG GAT GAA AAC TTC GAT CCT 1262
 Phe Phe Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro
 385 390 395 400

TTA CCT GAT TAT TGG CTA TCT CTT CTG TTC AAG AAA TTG GTG GGC ACC 1310
 Leu Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr
 405 410 415

AAG GTG TTA ATG GCA AGC GTG CAA GGT TCA AAG AGA AGG AAG CTT CGA 1358
 Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu Arg
 420 425 430

GTA TAC CTT CAT TGC ACA AAC ACT GAC AAT CCA AGG TAT AAA GAA GGA 1406
 Val Tyr Leu His Cys Thr Asn Thr Asp Asn Pro Arg Tyr Lys Glu Gly
 435 440 445

GAT TTA ACT CTG TAT GCC ATA AAC CTC CAT AAC GTC ACC AAG TAC TTG 1454
 Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr Lys Tyr Leu
 450 455 460

CGG TTA CCC TAT CCT TTT TCT AAC AAG CAA GTG GAT AAA TAC CTT CTA 1502
 Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp Lys Tyr Leu Leu
 465 470 475 480

AGA CCT TTG GGA CCT CAT GGA TTA CTT TCC AAA TCT GTC CAA CTC AAT 1550
 Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys Ser Val Gln Leu Asn
 485 490 495

GGT CTA ACT CTA AAG ATG GTG GAT GAT CAA ACC TTG CCA CCT TTA ATG 1598
 Gly Leu Thr Leu Lys Met Val Asp Asp Gln Thr Leu Pro Pro Leu Met
 500 505 510

GAA AAA CCT CTC CGG CCA GGA AGT TCA CTG GGC TTG CCA GCT TTC TCA 1646
 Glu Lys Pro Leu Arg Pro Gly Ser Ser Leu Gly Leu Pro Ala Phe Ser
 515 520 525

TAT AGT TTT TTT GTG ATA AGA AAT GCC AAA GTT GCT GCT TGC ATC TGA 1694
 Tyr Ser Phe Phe Val Ile Arg Asn Ala Lys Val Ala Ala Cys Ile
 530 535 540 543

AAA TAA AAT ATA CTA GTC CTG ACA CTG 1721

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 824
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12

CTGGCAAGAA GGTCTGGTTG GGAGAGACGA GCTCAGCTTA CGGTGGCGGT GCACCCTTGC 60
 TGTCCAACAC CTTTGCAGCT GGCTTTATGT GGCTGGATAA ATTGGGCCTG TCAGCCCAGA 120
 TGGGCATAGA AGTCGTGATG AGGCAGGTGT TCTTCGGAGC AGGCAACTAC CACTTAGTGG 180
 ATGAAAACCTT TGAGCCTTTA CCTGATTACT GGCTCTCTCT TCTGTTCAAG AAAGTGGTAG 240
 GTCCCAAGGT GTTACTGTCA AGAGTGAAAG GCCCAGACAG GAGCAAACCTC CGAGTGTATC 300
 TCCACTGCAC TAACGTCTAT CACCCACGAT ATCAGGAAGG AGATCTAACT CTGTATGTCC 360
 TGAACCTCCA TAATGTCACC AAGCACTTGA AGGTACCGCC TCCGTTGTTC AGGAAACCAG 420
 TGGATACGTA CCTTCTGAAG CCTTCGGGGC CGGATGGATT ACTTTCCAAA TCTGTCCAAC 480
 TGAACGGTCA AATTCTGAAG ATGGTGGATG AGCAGACCCT GCCAGCTTTG ACAGAAAAAC 540
 CTCTCCCCGC AGGAAGTGCA CTAAGCCTGC CTGCCTTTTC CTATGGTTTT TTTGTCATAA 600
 GAAATGCCAA AATCGCTGCT TGTATATGAA AATAAAAGGC ATACGGTACC CCTGAGACAA 660
 AAGCCGAGGG GGGTGTTATT CATAAAACAA AACCCTAGTT TAGGAGGCCA CCTCCTTGCC 720
 GAGTTCCAGA GCTTCGGGAG GTTGGGTGAC ACTTCAGTAT TACATTCACT GTGGTGTCT 780
 CTCTAAGAAG AATACTGCAG GTGGTGACAG TTAATAGCAC TGTG 824

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1899
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13

GGGAAAGCGA GCAAGGAAGT AGGAGAGAGC CGGGCAGGCG GGGCGGGGTT GGATTGGGAG 60
 CAGTGGGAGG GATGCAGAAG AGGAGTGGGA GGGATGGAGG GCGCAGTGGG AGGGGTGAGG 120
 AGGCGTAACG GGGCGGAGGA AAGGAGAAAA GGGCGCTGGG GCTCGGCGGG AGGAAGTGCT 180
 AGAGCTCTCG ACTCTCCGCT GCGCGGCAGC TGGCGGGGGG AGCAGCCAGG TGAGCCCAAG 240
 ATGCTGCTGC GCTCGAAGCC TGCGCTGCCG CCGCCGCTGA TGCTGCTGCT CCTGGGGCCG 300
 CTGGGTCCCT TCTCCCCTGG CGCCCTGCCG CGACCTGCGC AAGCACAGGA CGTCGTGGAC 360
 CTGGACTTCT TCACCCAGGA GCGCTGCAC CTGGTGAGCC CCTGTTTCTT GTCCGTCAAC 420
 ATTGACGCCA ACCTGGCCAC GGACCCGCGG TTCCTCATCC TCCTGGGTTT TCCAAAGCTT 480
 CGTACCTTGG CCAGAGGCTT GTCTCCTGCG TACCTGAGGT TTGGTGGCAC CAAGACAGAC 540
 TTCCTAATTT TCGATCCCAA GAAGGAATCA ACCTTTGAAG AGAGAAGTTA CTGGCAATCT 600
 CAAGTCAACC AGGATATTTG CAAATATGGA TCCATCCCTC CTGATGTGGA GGAGAAGTTA 660
 CGGTTGGAAT GGCCCTACCA GGAGCAATTG CTAATCCGAG AACACTACCA GAAAAAGTTC 720
 AAGAACAGCA CCTACTCAAG AAGCTCTGTA GATGTGCTAT AACTTTTGC AAAGTGTCTA 780
 GGACTGGACT TGATCTTTGG CCTAAATGCG TTATTAAGAA CAGCAGATTT GCAGTGAAC 840
 AGTTCTAATG CTCAGTTGCT CTGGACTAC TGCTCTTCCA AGGGGTATAA CATTTCTTGG 900


```
(2)      INFORMATION FOR SEQ ID NO:14:
      (1)      SEQUENCE CHARACTERISTICS:
                (A)      LENGTH:          592
                (B)      TYPE:             amino acid
                (C)      STRANDEDNESS:     singl
                (D)      TOPOLOGY:        linear
      (xi)     SEQUENCE DESCRIPTION:  SEQ ID NO:14
```

Met	Glu	Gly	Ala	Val	Gly	Gly	Val	Arg	Arg	Arg	Asn	Gly	Ala	Glu
				5					10					15
Glu	Arg	Arg	Lys	Gly	Arg	Trp	Gly	Ser	Ala	Gly	Gly	Ser	Ala	Arg
				20					25					30
Ala	Leu	Asp	Ser	Pro	Leu	Arg	Gly	Ser	Trp	Arg	Gly	Glu	Gln	Pro
				35					40					45
Gly	Glu	Pro	Lys	Met	Leu	Leu	Arg	Ser	Lys	Pro	Ala	Leu	Pro	Pro
				50					55					60
Pro	Leu	Met	Leu	Leu	Leu	Leu	Gly	Pro	Leu	Gly	Pro	Leu	Ser	Pro
				65					70					75
Gly	Ala	Leu	Pro	Arg	Pro	Ala	Gln	Ala	Gln	Asp	Val	Val	Asp	Leu
				80					85					90
Asp	Phe	Phe	Thr	Gln	Glu	Pro	Leu	His	Leu	Val	Ser	Pro	Ser	Phe
				95					100					105
Leu	Ser	Val	Thr	Ile	Asp	Ala	Asn	Leu	Ala	Thr	Asp	Pro	Arg	Phe
				110					115					120
Leu	Ile	Leu	Leu	Gly	Ser	Pro	Lys	Leu	Arg	Thr	Leu	Ala	Arg	Gly
				125					130					135
Leu	Ser	Pro	Ala	Tyr	Leu	Arg	Phe	Gly	Gly	Thr	Lys	Thr	Asp	Phe
				140					145					150
Leu	Ile	Phe	Asp	Pro	Lys	Lys	Glu	Ser	Thr	Phe	Glu	Glu	Arg	Ser
				155					160					165
Tyr	Trp	Gln	Ser	Gln	Val	Asn	Gln	Asp	Ile	Cys	Lys	Tyr	Gly	Ser
				170					175					180
Ile	Pro	Pro	Asp	Val	Glu	Glu	Lys	Leu	Arg	Leu	Glu	Trp	Pro	Tyr
				185					190					195
Gln	Glu	Gln	Leu	Leu	Leu	Arg	Glu	His	Tyr	Gln	Lys	Lys	Phe	Lys
				200					205					210
Asn	Ser	Thr	Tyr	Ser	Arg	Ser	Ser	Val	Asp	Val	Leu	Tyr	Thr	Phe
				215					220					225
Ala	Asn	Cys	Ser	Gly	Leu	Asp	Leu	Ile	Phe	Gly	Leu	Asn	Ala	Leu

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1899

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15

															GGG	3
AAA	GCG	AGC	AAG	GAA	GTA	GGA	GAG	AGC	CGG	GCA	GGC	GGG	GCG	GGG		48
TTG	GAT	TGG	GAG	CAG	TGG	GAG	GGA	TGC	AGA	AGA	GGA	GTG	GGA	GGG		93
ATG	GAG	GGC	GCA	GTG	GGA	GGG	GTG	AGG	AGG	CGT	AAC	GGG	GCG	GAG		138
Met	Glu	Gly	Ala	Val	Gly	Gly	Val	Arg	Arg	Arg	Asn	Gly	Ala	Glu		
				5					10					15		
GAA	AGG	AGA	AAA	GGG	CGC	TGG	GGC	TCG	GCG	GGA	GGA	AGT	GCT	AGA		183
Glu	Arg	Arg	Lys	Gly	Arg	Trp	Gly	Ser	Ala	Gly	Gly	Ser	Ala	Arg		
				20					25					30		
GCT	CTC	GAC	TCT	CCG	CTG	CGC	GGC	AGC	TGG	CGG	GGG	GAG	CAG	CCA		228
Ala	Leu	Asp	Ser	Pro	Leu	Arg	Gly	Ser	Trp	Arg	Gly	Glu	Gln	Pro		
				35					40					45		
GGT	GAG	CCC	AAG	ATG	CTG	CTG	CGC	TCG	AAG	CCT	GCG	CTG	CCG	CCG		273
Gly	Glu	Pro	Lys	Met	Leu	Leu	Arg	Ser	Lys	Pro	Ala	Leu	Pro	Pro		
				50					55					60		
CCG	CTG	ATG	CTG	CTG	CTC	CTG	GGG	CCG	CTG	GGT	CCC	CTC	TCC	CCT		318
Pro	Leu	Met	Leu	Leu	Leu	Leu	Gly	Pro	Leu	Gly	Pro	Leu	Ser	Pro		
				65					70					75		
GGC	GCC	CTG	CCC	CGA	CCT	GCG	CAA	GCA	CAG	GAC	GTc	GTG	GAC	CTG		363
Gly	Ala	Leu	Pro	Arg	Pro	Ala	Gln	Ala	Gln	Asp	Val	Val	Asp	Leu		
				80					85					90		
GAC	TTC	TTC	ACC	CAG	GAG	CCG	CTG	CAC	CTG	GTG	AGC	CCC	TCG	TTC		408
Asp	Phe	Phe	Thr	Gln	Glu	Pro	Leu	His	Leu	Val	Ser	Pro	Ser	Phe		
				95					100					105		
CTG	TCC	GTC	ACC	ATT	GAC	GCC	AAC	CTG	GCC	ACG	GAC	CCG	CGG	TTC		453
Leu	Ser	Val	Thr	Ile	Asp	Ala	Asn	Leu	Ala	Thr	Asp	Pro	Arg	Phe		
				110					115					120		
CTC	ATC	CTC	CTG	GGT	TCT	CCA	AAG	CTT	CGT	ACC	TTG	GCC	AGA	GGC		498
Leu	Ile	Leu	Leu	Gly	Ser	Pro	Lys	Leu	Arg	Thr	Leu	Ala	Arg	Gly		
				125					130					135		
TTG	TCT	CCT	GCG	TAC	CTG	AGG	TTT	GGT	GGC	ACC	AAG	ACA	GAC	TTC		543
Leu	Ser	Pro	Ala	Tyr	Leu	Arg	Phe	Gly	Gly	Thr	Lys	Thr	Asp	Phe		
				140					145					150		
CTA	ATT	TTC	GAT	CCC	AAG	AAG	GAA	T								

Gln	Glu	Gln	Leu	Leu	Leu	Arg	Glu	His	Tyr	Gln	Lys	Lys	Phe	Lys	
				200					205					210	
AAC	AGC	ACC	TAC	TCA	AGA	AGC	TCT	GTA	GAT	GTG	CTA	TAC	ACT	TTT	768
Asn	Ser	Thr	Tyr	Ser	Arg	Ser	Ser	Val	Asp	Val	Leu	Tyr	Thr	Phe	
				215					220					225	
GCA	AAC	TGC	TCA	GGA	CTG	GAC	TTG	ATC	TTT	GGC	CTA	AAT	GCG	TTA	813
Ala	Asn	Cys	Ser	Gly	Leu	Asp	Leu	Ile	Phe	Gly	Leu	Asn	Ala	Leu	
				230					235					240	
TTA	AGA	ACA	GCA	GAT	TTG	CAG	TGG	AAC	AGT	TCT	AAT	GCT	CAG	TTG	858
Leu	Arg	Thr	Ala	Asp	Leu	Gln	Trp	Asn	Ser	Ser	Asn	Ala	Gln	Leu	
				245					250					255	
CTC	CTG	GAC	TAC	TGC	TCT	TCC	AAG	GGG	TAT	AAC	ATT	TCT	TGG	GAA	903
Leu	Leu	Asp	Tyr	Cys	Ser	Ser	Lys	Gly	Tyr	Asn	Ile	Ser	Trp	Glu	
				260					265					270	
CTA	GGC	AAT	GAA	CCT	AAC	AGT	TTC	CTT	AAG	AAG	GCT	GAT	ATT	TTC	948
Leu	Gly	Asn	Glu	Pro	Asn	Ser	Phe	Leu	Lys	Lys	Ala	Asp	Ile	Phe	
				275					280					285	
ATC	AAT	GGG	TCG	CAG	TTA	GGA	GAA	GAT	TAT	ATT	CAA	TTG	CAT	AAA	993
Ile	Asn	Gly	Ser	Gln	Leu	Gly	Glu	Asp	Tyr	Ile	Gln	Leu	His	Lys	
				290					295					300	
CTT	CTA	AGA	AAG	TCC	ACC	TTC	AAA	AAT	GCA	AAA	CTC	TAT	GGT	CCT	1038
Leu	Leu	Arg	Lys	Ser	Thr	Phe	Lys	Asn	Ala	Lys	Leu	Tyr	Gly	Pro	
				305					310					315	
GAT	GTT	GGT	CAG	CCT	CGA	AGA	AAG	ACG	GCT	AAG	ATG	CTG	AAG	AGC	1083
Asp	Val	Gly	Gln	Pro	Arg	Arg	Lys	Thr	Ala	Lys	Met	Leu	Lys	Ser	
				320					325					330	
TTC	CTG	AAG	GCT	GGT	GGA	GAA	GTG	ATT	GAT	TCA	GTT	ACA	TGG	CAT	1128
Phe	Leu	Lys	Ala	Gly	Gly	Glu	Val	Ile	Asp	Ser	Val	Thr	Trp	His	
				335					340					345	
CAC	TAC	TAT	TTG	AAT	GGA	CGG	ACT	GCT	ACC	AGG	GAA	GAT	TTT	CTA	1173
His	Tyr	Tyr	Leu	Asn	Gly	Arg	Thr	Ala	Thr	Arg	Glu	Asp	Phe	Leu	
				350					355					360	
AAC	CCT	GAT	GTA	TTG	GAC	ATT	TTT	ATT	TCA	TCT	GTG	CAA	AAA	GTT	1218
Asn	Pro	Asp	Val	Leu	Asp	Ile	Phe	Ile	Ser	Ser	Val	Gln	Lys	Val	
				365					370					375	
TTC	CAG	GTG	GTT	GAG	AGC	ACC	AGG	CCT	GGC	AAG	AAG	GTC	TGG	TTA	1263
Phe	Gln	Val	Val	Glu	Ser	Thr	Arg	Pro	Gly	Lys	Lys	Val	Trp	Leu	
				380					385					390	
GGA	GAA	ACA	AGC	TCT	GCA	TAT	GGA	GGC	GGA	GCG	CCC	TTG	CTA	TCC	1308
Gly	Glu	Thr	Ser	Ser	Ala	Tyr	Gly	Gly	Gly	Ala	Pro	Leu	Leu	Ser	
				395					400					405	
GAC	ACC	TTT	GCA	GCT	GGC	TTT	ATG	TGG	CTG	GAT	AAA	TTG	GGC	CTG	1353
Asp	Thr	Phe	Ala	Ala											

410	415	420	
TCA GCC CGA ATG GGA ATA gAA GTG GTG ATG AGG CAA GTA TTC TTT			1398
Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val Phe Phe			
425	430	435	
GGA GCA GGA AAC TAC CAT TTA GTG GAT GAA AAC TTC GAT CCT TTA			1443
Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro Leu			
440	445	450	
CCT GAT TAT TGG CTA TCT CTT CTG TTC AAG AAA TTG GTG GGC ACC			1488
Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr			
455	460	465	
AAG GTG TTA ATG GCA AGC GTG CAA GGT TCA AAG AGA AGG AAG CTT			1533
Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu			
470	475	480	
CGA GTA TAC CTT CAT TGC ACA AAC ACT GAC AAT CCA AGG TAT AAA			1578
Arg Val Tyr Leu His Cys Thr Asn Thr Asp Asn Pro Arg Tyr Lys			
485	490	495	
GAA GGA GAT TTA ACT CTG TAT GCC ATA AAC CTC CAT AAC GTC ACC			1623
Glu Gly Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr			
500	505	510	
AAG TAC TTG CGG TTA CCC TAT CCT TTT TCT AAC AAG CAA GTG GAT			1668
Lys Tyr Leu Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp			
515	520	525	
AAA TAC CTT CTA AGA CCT TTG GGA CCT CAT GGA TTA CTT TCC AAA			1713
Lys Tyr Leu Leu Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys			
530	535	540	
TCT GTC CAA CTC AAT GGT CTA ACT CTA AAG ATG GTG GAT GAT CAA			1758
Ser Val Gln Leu Asn Gly Leu Thr Leu Lys Met Val Asp Asp Gln			
545	550	555	
ACC TTG CCA CCT TTA ATG GAA AAA CCT CTC CGG CCA GGA AGT TCA			1803
Thr Leu Pro Pro Leu Met Glu Lys Pro Leu Arg Pro Gly Ser Ser			
560	565	570	
CTG GGC TTG CCA GCT TTC TCA TAT AGT TTT TTT GTG ATA AGA AAT			1848
Leu Gly Leu Pro Ala Phe Ser Tyr Ser Phe Phe Val Ile Arg Asn			
575	580	585	
GCC AAA GTT GCT GCT TGC ATC TGA AAA TAA AAT ATA CTA GTC CTG			1893
Ala Lys Val Ala Ala Cys Ile			
590	592		
ACA CTG			1899

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 594
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16

```
ATTACTATAG GGCACGCGTG GTCGACGGCC CGGGCTGGTA TTGTCTTAAT GAGAAGTTGA 60
TAAAGAATTT TGGGTGGTTG ATCTCTTTCC AGCTGCAGTT TAGCGTATGC TGAGGCCAGA 120
TTTTTTCAGG CAAAAGTAAA ATACCTGAGA AACTGCCTGG CCAGAGGACA ATCAGATTTT 180
GGCTGGCTCA AGTGACAAGC AAGTGTTTAT AAGCTAGATG GGAGAGGAAG GGATGAATAC 240
TCCATTGGAG GCTTTACTCG AGGGTCAGAG GGATACCCGG CGCCATCAGA ATGGGATCTG 300
GGAGTCGGAA ACGCTGGGTT CCCACGAGAG CGCGCAGAAC ACGTGCCTCA GGAAGCCTGG 360
TCCGGGATGC CCAGCGCTGC TCCCCGGGCG CTCCTCCCCG GCGCTCCTC CCCAGGCCTC 420
CCGGGCGCTT GGATCCCGGC CATCTCCGCA CCCTTCAAGT GGGTGTGGGT GATTTCGTAA 480
GTGAACGTGA CCGCCACCGG GGGGAAAGCG AGCAAGGAAG TAGGAGAGAG CCGGGCAGGC 540
GGGGCGGGT TGGATTGGGA GCAGTGGGAG GGATGCAGAA GAGGAGTGGG AGGG 594
```

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17
CCCCAGGAGC AGCAGCATCA G 21

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18
AGGCTTCGAG CGCAGCAGCA T 21

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19
GTAATACGAC TCACTATAGG GC 22

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20
ACTATAGGC ACGCGTGGT 19

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21
CTTGGGCTCA CCTGGCTGCT C 21

- (2) INFORMATION FOR SEQ ID NO:22:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22
AGCTCTGTAG ATGTGCTATA CAC 23
- (2) INFORMATION FOR SEQ ID NO:23:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23
GCATCTTAGC CGTCTTTCTT CG 22
- (2) INFORMATION FOR SEQ ID NO:24:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24
GAGCAGCCAG GTGAGCCCAA GAT 23
- (2) INFORMATION FOR SEQ ID NO:25:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25
TTCGATCCCA AGAAGGAATC AAC 23
- (2) INFORMATION FOR SEQ ID NO:26:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26
AGCTCTGTAG ATGTGCTATA CAC 23
- (2) INFORMATION FOR SEQ ID NO:27:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27
TCAGATGCAA GCAGCAACTT TGGC 24
- (2) INFORMATION FOR SEQ ID NO:28:
- (i) SEQUENCE CHARACTERISTICS:

0998813 11904

(A) LENGTH: 22
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28
GCATCTTAGC CGTCTTTCTT CG 22

(2) INFORMATION FOR SEQ ID NO:29:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29
GTAGTGATGC CATGTAACATG AATC 24

(2) INFORMATION FOR SEQ ID NO:30:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30
AGGCACCCTA GAGATGTTCC AG 22

(2) INFORMATION FOR SEQ ID NO:31:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31
GAAGATTCTT GTTCCATGA CGTG 24

(2) INFORMATION FOR SEQ ID NO:32:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32
CCACACTGAA TGTAATACTG AAGTG 25

(2) INFORMATION FOR SEQ ID NO:33:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33
CGAAGCTCTG GAACTCGGCA AG 22

(2) INFORMATION FOR SEQ ID NO:34:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34
 GCCAGCTGCA AAGGTGTTGG AC 22

(2) INFORMATION FOR SEQ ID NO:35:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 23
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35
 AACACCTGCC TCATCAGAC TTC 23

(2) INFORMATION FOR SEQ ID NO:36:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36
 GCCAGGCTGG CGTCGATGGT GA 22

(2) INFORMATION FOR SEQ ID NO:37:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37
 GTCGATGGTG ATGGACAGGA AC 22

(2) INFORMATION FOR SEQ ID NO:38:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38
 GTAATACGAC TCACTATAGG GC 22

(2) INFORMATION FOR SEQ ID NO:39:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 19
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39
 ACTATAGGGC ACGCGTGGT 19

(2) INFORMATION FOR SEQ ID NO:40:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 27
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40
 CCATCCTAAT ACGACTCACT ATAGGGC 27

095813 11901

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41

ACTCACTATA GGGCTCGAGC GGC 23

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44848
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42

GGATCTTGGC TCACTGCAAT CTCTGCCTCC CATGCAATTC TTATGCATCA 50
 GCCTCCTGAG TAGCTTGGAT TATAGGCTCG CGCCACCACT CCTGGCTACA 100
 CCATGTTGCC CAGGCTGGTC TTGAACCTTT GGGCTCTAGT GATCCACCCG 150
 CCTTGGCCTC CCAGAGTGCT GGGATTACAG GTGTGAGCCA TCACACCCGG 200
 CCCCCCGTTT CCATATTAGT AACTCACATG TAGACCACAA GGATGCACTA 250
 TTTAGAAAAC TTGCAATGGT CCACTTTTCA AATCACCACAA ACATGTTAAA 300
 GAAATTGGTA TGAATGGGCA TGGCACAGTG GCTCATGCCT GCAATCCTAG 350
 CATTGTGTGA GGGTGAGACG GGCAGATCAC GAGGTGAGGA GATTGAGACC 400
 ATCCTGACAG ACATGGTGAA ATCCCATCTC TACTAAAAAT ACAAAACAAT 450
 TAGCCGGGGG TGATGGCAGG CCCCTGTAGT CCCAGCTACT CGGGAGGCTG 500
 AGGCAGGAGA ATGGCGTGAA TCCAGGAGGC AGAGCTTGCA GTGAGCCGAG 550
 ATGGTGCCAC TGCACCTCCAG CCTGGGCGAC AGAGCGAGAC TCCGTCTCAA 600
 AAAAAAAAAA AAAGAAAGAA ATTGGTATGA CTGTTGACTC ACAACAGGAG 650
 TCAGGGGCAAT GGGGTGGGGT GTAAGATTAA TGTGATGACA AATGTGGAAA 700
 AGAAACTTCT GTTTTTCCTA CTCCACGTCT GCTACCATA TATTACACTC 750
 TTCTGGTAGT GTGGTGTGTTA TGTGTGAATT TTTTTCATA TGTATACAGT 800
 AATTGTAGGA TATGAACCTG ATTCTAGTTG CAAAACCTAC TATGAGCTTA 850
 GCTTTTAAGT TGCTTAAGAA TAGGTAGATC TATGCAAATA ATGATAATTA 900
 TTATTATTAT TTTAAGAGAG GGTCTCACCT TGTACCCAG GCTGGAGTGC 950
 AGTGGTGTGA TTAAGGGTCA CTGCAACCTC CACCTCCCAG GCTCAAATAA 1000
 ACCTCCCAC TCAGCCTCCC CAGTAGCTGG AACCACAGGC ACGGGCCACC 1050
 ACGCCTGGCT AATTTTTTGT ATTTTTTGTG GAGATGGGGT TTCATCATGT 1100
 TGCCCGAGCT GTTCTTGAAT TCCTCGGCTC AAGCAATCCT CCCACCTTGG 1150
 CCTCCCAAA TGCTGGCATC ACAGGCATGA TGGCATCACT GGCATCACAT 1200
 ACCATGCCTG GCCTGATTTA TGCAAATTAG ATATGCATT CAAAATAATC 1250
 TATTTTTATT TGTTGCCTTA TTGGTGGTAC AATCTCAAGT GGAAAAATCT 1300
 AAGGGTTTTG GTGTTATTG CTTACTCAAC CAATATTAT TAGACTCTTA 1350
 CTAAGCACCA ACATGATCAC ATGCCTGAGC TATGGCTAGC ATAGCGTGTG 1400
 AGACAAACTT AATCTCTGTT TTGGTGGAGC ATATAATCTA GTAGATGAAG 1450
 CCAATGTTGA GCAACATCAC AATACTAACA AATTGAGGAT GCTACGAGAG 1500
 TGTCTAACAA ATTGAGGATG CTACGAGAGT GTCTAACAAA TTGAGGATGC 1550
 TATGAGAGTG TGTGATGGAG AGCTGCCTGG AGATTGAGAG AAAGCTTCCT 1600
 TGAGGGAAGT TACATTTCAG CTGAAACACA CTGCCATCTG CTCGAGGTTT 1650
 TGTAACTGCA TTCACATCCC GATTCTGACA CTTACATCC CGATTCTGAC 1700
 ACTTCACCCA GTTACTGTCT CAGAGCTTGG GTCCGCATGT GTAAAAAAG 1750
 GACAGTATGC ACTTGGCAGG GTTGTGAGAA GGAAGAGAAA CACAAGTAAA 1800
 GCACCTGTAT CAGGCATACA GTAGGCACTA AGCGTGCAT GCTTGCTATG 1850
 ATTATACATC AGTGTAAAGC TCAAGGAAAA GCTGAAGAAA AGTCTGACCA 1900
 ACAGCGAAG ATAAATGCGC AGAGGAGAAA TTTGGCAAAG GCTCCAAATT 1950
 CAGGGGCAGT CCGTACTCTA CACTTTGTAT GGGGGCTTCA GGTCTCTGAGT 2000
 TCCAGACATT GGAGCACTA ACCCTTAAAG ATTGCTAAAT ATTGTCTTAA 2050
 TGAGAAGTTG ATAAAGAATT TTGGGTGGTT GATCTCTTTC CAGCTGCAGT 2100
 TTAGCGTATG CTGAGGCCAG ATTTTTTCAA GCAAAGTAA AATACCTGAG 2150
 AAATGCGCTG GCCAGAGGAC AATCAGATTT TGGCTGGCTC AAGTGACAAG 2200
 CAAGTGTTTA TAAGTAGAT GGGAGAGGAA GGGATGAATA CTCCATGGA 2250
 GGTTTTACTC GAGGGTCAAG GGGATACCCG GCGCCATCAG AATGGGATCT 2300
 GGGAGTCGGA AACGCTGGGT TCCCACGAGA GCGCGCAGAA CACGTGCGTC 2350
 AGGAAGCCTG GTCCGGGATG CCCAGCGCTG CTCCCGGGG GCTCCTCCC 2400
 GGGCGCTCCT CCCAGGCCCT CCCGGGCGCT TGGATCCCGG CCATCTCCGC 2450
 ACCCTTCAAG TGGGTGTGGG TGATTTCGTA AGTGAACGTG ACCGCCACCG 2500
 AGGGGAAAGC GAGCAAGGAA GTAGGAGAGA GCCGGCAGG CGGGCGGGG 2550
 TTGGATTGGG AGCAGTGGGA GGGATGCAGA AGAGGAGTGG GAGGGATGGA 2600
 GGGCGCAGTG GGAGGGGTGA GGAGGCGTAA CGGGGCGGAG GAAAGGAGAA 2650
 AAGGCGCTG GGGCTCGGCG GGAGGAAGTG CTAGAGCTCT CGACTCTCCG 2700
 CTGCGCGGCA GCTGGCGGGG GGAGCAGCCA GGTGAGCCCA AGATGCTGCT 2750

GCGCTCGAAG	CCTGCGCTGC	CGCCGCCGCT	GATGCTGCTG	CTCCTGGGGC	2800
CGCTGGGTCC	CCTCTCCCCT	GGCGCCCTGC	CCCGACCTGC	GCAAGCACAG	2850
GACGTCGTGG	ACCTGGACCT	CTTCACCCAG	GAGCCGCTGC	ACCTGGTGAG	2900
CCCTCGTCTC	CTGTCCGTCA	CCATTGACGC	CAACCTGGCC	ACGGACCCGC	2950
GGTTCCCTCAT	CCTCCTGGGG	TAAGCGCCAG	CCTCCTGGTC	CTGTCCCTTT	3000
TCCTGTCTCT	CTGACACCTA	TGCTGCCCC	GCCAGCGGCT	CTCCTTCTTT	3050
TGCGCGGAAA	CAACTTCACA	CCGGAACCTC	CCCGCCTGTC	TCTCCCCACC	3100
CCACTTCCCG	CCTCTCATT	TCCCTCTCCC	TCCCTTACTC	TCAGACCCCA	3150
AACCGCTTTT	TGGGGGGTAT	CATTTAAAA	ATAGATTTAG	GGGTACAAAG	3200
TGCAGTTCTG	TTCCATGGGT	ATATTGCATT	GTGGTGGCAT	CTGGGCTCTT	3250
AGTGTAACCTG	TCACCCGAAT	GTTGTACATT	GTATCTAATA	GGTAATTCTT	3300
CATCCCTCAT	CCCTCTCCCA	CCCTCCCACC	TTTTGGAGTC	TCCAGTGTCT	3350
ACTATTCCAC	TAAGTCCATG	TGTACACATT	GTTTAGCGCC	CACCTCTAAAT	3400
GAGCCTTTTT	GTTTCATTCA	TTCTGTAAGT	GTTGAATAGG	CACCACCTAA	3450
GGTCAGGTAT	AAGTGAAAT	TTGAAAAAGA	AACTGCCAC	TTGCCCCAGT	3500
ACTTCCCTAG	CCAAGAGGAG	GGAACCAGG	CAGGTGCACC	TGAAGGCTG	3550
TGAGTGCTTG	ATTTGCTGTG	CAGTGATAGG	CAAGTAAGAT	TGTGCATAGC	3600
CTTCTGTATT	TAAGACTGTG	TTAGGAAGAT	TTCTCTTTCT	TTTCTTTTCT	3650
TTTTCTTTTT	TCTTTTCTTT	TTTTTTTTTA	GGCAGATGAA	AAGGGCTCTA	3700
CAGAACAGGA	ATAAAAAATCT	AAATATTCAA	TAAATGAGAC	CTAGGAGACT	3750
ACTGCAGTGA	CTTACAAAGT	CCTAATAAAA	AGATGTCTCT	CCAAAAATGG	3800
GCTGCAAAAT	TGGGTGCCATG	CTTATCAGCT	CTAAGTTTTT	TCCTTACCTG	3850
AGAAAGAGG	AACCTGATGC	AGGTTCCAGG	CTCCTGCCCC	ATGAATGCAG	3900
GCTGACTCCA	AGATGGGGAG	CTACAGGGAC	AAATCCAGGT	CTTCTAGGCC	3950
TCTTATTTAG	GCCCTGGGAG	CCTCCAGAGA	TGGCCACATC	TTGACCAGCC	4000
CAGATAGAGG	GAAAGATCAC	CATTATCTCA	CCTCTGTGTC	AAATACCTAG	4050
ATGCTGTCTT	CCCTGAGCCC	ACACTATAGT	TGCCAGCGCT	AATTTAATGG	4100
GTAGTGTACT	GGTTAAGAGA	TGGACAGACC	ATCCTGGCTT	GACTCTCAGC	4150
TCTGGCAAA	ATGAGTGACT	TGGTTTTTCC	ATATCTCTTG	GCCACACCAA	4200
CCCTGATTTT	TTACAGCTGA	GAATGGAATT	TCTCAAGCTT	GCCTCAAGGA	4250
TTATTGGCCG	AGGATTTGAT	GATATGGTAA	GAGCTTCTCA	GTGTTTGACC	4300
CATAGTAAGT	GTTTGACGTT	TCAAACGAAT	TGTTTCTTTC	TAGGACATGG	4350
TGAGCATTGG	GTAGCCATTC	ACCGGTTTTT	TGTTTCTTTG	GATCATAGTT	4400
AACTCTCTCT	TTTCTCTCTG	GCACTACAAT	TTTCTGGTGG	GGAAGAATCC	4450
TTACTTTCTG	CCCTTCCCCT	TAAGGATAGG	AAGCTGATAC	TAGGCAGCAA	4500
CTAGTTGGGG	GATAGGAAGA	TTGTTCCAGA	GAAATGCTGA	ACCATAGGGC	4550
TCCAGATCAC	AGGACCCAG	TCTTAGCTTG	CTGGGGTGTG	GGGTGGGGGG	4600
GGGCGGTAC	TGAACATGGG	TATGAAGTAG	ATGTCCATTT	ACTGAAATGT	4650
GAGGACCTGA	GGCCTCTTCT	ATTGCTGTAG	CCAGCATATT	CCCCAACCTC	4700
TCCCAAGAAA	AGGACAGATG	GGGGTTCCCC	CCTGGAGTAA	CAGGTCCAAA	4750
AGAAAAAACA	TACAGTGGGA	CTTCCAGGAT	CTGGGCCTGA	TCACCCAGCA	4800
GTCAGGCTCC	CCGCAATTGA	CTAACACCCC	CCTAACACGT	AGAAATTCCA	4850
ATCTGCAATT	TAGTGAGGAT	GATACCTTTA	TTCTTCTTAA	ATACATCTCT	4900
TCATTTCCCA	GAGCACCCCT	TTTTCCCTCT	CTCTGCACCT	TTTTGTAAAA	4950
GACTGGAGTA	TAATGAAATA	CCAAGAGAGC	ATAACATGTG	ATACATAAAA	5000
CTTTTTTTCT	GGTTTACAAA	ACAGTTCATT	CTTGTCCTA	CGTGCTTCTC	5050
TCCAAGGCTG	GCTGCTGTCT	GTTCAGCCCC	GCTTCGCTTG	GAGAGGCCAT	5100
CTGCCATACC	TGCTCCCCAG	ACGCATCGAC	AAGCACACCC	AGAGTGTAT	5150
CTGCTAAGAC	CTAAAAGAGG	GAGGAACCCC	CTCTCTCAT	CTAAGACCTA	5200
GCTTCTAAAT	TAGAGTGTGA	GGGTCCATCT	CCCCAGGAGG	GGCACAGGGC	5250
CCAAACAGCC	CAGCCATCTC	AGAAGACAAC	ACTAAGCTTT	GTAGGGGTCC	5300
ACAGTAGAGG	AGAGTAAGAC	GCCTGTTGTT	TAATTTATTA	CAGTTCCTCA	5350
AAAGTGAAAG	TGTGTGGGCG	GGATGGCAAG	AGCTGAGCAG	ACGAAAGCTG	5400
AAGGAATTAAG	GAAAGAGAGG	AGGACACAAA	CAGCTGACAC	TTCTCTAGTT	5450
CTTGTCATTT	GCCTGGCCCT	GTCTAAGCA	CCTTCTAGGT	ATTAATCCAT	5500
TTAGTCTTGG	CTACAACT	GTGAGTAACT	AGTTTGTCTA	CCCCATTTT	5550
AAAAATGAAG	AAAGTGAGGC	TCAGGGAGGT	TAAAGTAACT	GGCCACAGTT	5600
TGAAACTAGA	CTCTGATCAC	ATGAGATAAT	AGTGCCCAT	AAAAGGAAAA	5650
GCAGATTATA	TTTTTTAAAG	GAAAGAGAGT	AGGATATGGT	AGAAAAAGAT	5700
TGTTTGGAAA	GGAATTGAGA	GATTGATATA	ATGAAAAGAA	GCATTACAT	5750
GAGAGTAACA	GTATCAGGGC	CCAAACCTTC	ATCTAAGGTA	CTTCAAAGAG	5800
GCCTAAGCAA	ACTTAGTCTC	TGGCGTGGTT	CTAGTCTCCA	TGATGGCAAA	5850
TACATTGTGT	ACAGCCCAAC	TCCACACAAA	ACTTAAATAC	CAATGATAGA	5900
GCAATCTAAA	ATTTGAAAGA	AAAAATCTTT	CAATTTGTCT	TCTTCCCAGA	5950
GGGACTTAAT	CAAGAAACCA	ATCAAAATAC	TTCTAAGCC	TAAGTGTGTG	6000
CAGAACTCCA	AAGAGAGCCC	AGCCCTAAAT	CAACACTGTC	CAATGGAAAT	6050
ATAATATAAT	GTGGGCCTCA	TATGCAAGGT	CATATGTAAT	TTTAAATTTT	6100
CTAGTAGCCA	TATTAATAAG	GTAAAAAGAA	ACAAGTGAAA	TTAATTTTAA	6150
TAATTTTATT	TAGTTCAATA	GATCCAAAAT	GTTTCTCTAG	CATGTAATCA	6200
ATATAAAAAT	ATTAATGAGG	TATTTATTAT	TCCTTTTCTC	AAACCAAGTC	6250
TATTTCTATA	TCTGGCGTGT	ATTATTTACA	GCACCTCTCA	GACTATATTT	6300
CTTCTTTTCT	TTTTTTTTTC	CGAGACAATT	TTGCTCTTGT	CACCCAGCT	6350
AGAGTACAAT	GGCGTTACCT	CGGCTCACTG	CAACCTCCGC	CTCCCGGTT	6400
CAAGTTATTC	TCCTGCCCTA	GTCTCCCAAG	TAGCTGGGAC	TAGAGGCATG	6450
CACCACCAGC	CCTGCTAAT	TGTGTATTTT	TAGTAGAGAC	AGGGTTTCAC	6500
CATGTTGGCC	AGGCTAATCT	CAAACCTCTG	AGCTCAGGTG	ATATGCCAC	6550
CTCGCCCTTC	CAAAGTGTG	GGATTACAGG	CGTGAGCCAC	TGACCCCGGC	6600
CTCAGATTAA	CTATATTTCA	AGCGTTTCA	AGCCACATGT	AGCTAGTGCT	6650
ATGGTAGTGG	ACAGTACAGA	TCTGCATTTT	AATTAAGACA	CGTATACAAG	6700
CATAGTTTAC	TAATGCACGG	TAAAAAAGG	TATAGTGCTG	AGTCGGTGGT	6750

AGAAATCCTA	AATACTGCAG	AGCAAAAGTG	GTACGAACAG	CAATCTCAGT	6800
GATAATGCAA	CCATGCTTGC	TTTTCATTCG	AATTTGCTTA	TTTTCTTCA	6850
GCAAAGTTCA	TCCATTTTTG	CCAATTCAT	AAATATTTAC	TGATAAAAC	6900
TTTCAATATT	AGATTCTTGC	ATCTTCATAG	ACAGAGTTGC	TTTTACATT	6950
TAGAAAATTA	CTTATCAATG	TTAAACACAC	GTTTTGATAA	CCAGTGTGG	7000
AAAGAGGTGC	AGACTCCCCA	TGTGCCTATT	GATGGCAGAA	ATATTCACAG	7050
CCAAAGGGAA	ACAAAGGGCT	GGGACAAATC	ACACACCTCA	TGCTCCCTAA	7100
CTCCTGGGAA	GTGCTGTCCC	TCTGATTGAG	CTCTTATTAT	TGCCTTCCCC	7150
ACTAACCTTG	TCCACTGTGC	CCTGGAGCCC	TTTGCAGGGT	TACCTGTCTCT	7200
GTCCTCCTCA	CAGAATATCT	CCTCTACCTC	CTTGTCCTAAG	CTACAACTTG	7250
GCTATTCTCT	GATGACACTG	TCTTCCCCTG	AGCCCTTTTG	AGTAATGGCT	7300
GCATATTCTC	CCATAGTCCA	GTCTTTTTC	TGTTCTCCAG	TCTGGCTTCT	7350
GGATGACAGC	CCACTAGTTT	GAACCTCATA	CTGCTATAGT	TCAAGTCCCT	7400
TTTGACTTGT	TACCTTGGGC	AAATTACCTC	CTTTTGTTC	GGTTCCTTGT	7450
TTGTAAATG	ACGATAATAA	TGCCATTTCG	TTTCACTGGG	TATTTTGAAA	7500
TTGAGTGAAA	GAAGGGGGT	AGCTTCCCTA	CACGCTCAGT	GTAGACTAGC	7550
CTGATGTGCA	TTACGGGTGA	TGCCATGACT	CAGTGTGTTT	TCCTCATCTC	7600
CACATCTGGC	TCTCATCCAG	TGCTCCTGCT	TACGGCACTC	TGTCCTCTCT	7650
TTACTTACTC	CCCCATTATTA	ACTGAAGACT	GGCACTGATC	TCACAGTTTC	7700
CTCTCCACTT	CCTAGTCTCA	CCATCATCCT	AGATGACTTC	AAGTCACCTA	7750
GATAAACTGT	CTCAGTTTCT	TCACTCACAT	TTTTTTATAA	CAGATAATGT	7800
TACACTCAAG	TTGTAACAGA	ACCAGCTTAT	CCAGCTCATG	AAATGTATGC	7850
ATTTTCATCT	AACCTGTGAT	TCAGTGACAT	CCTGTGGGTA	TCTGGAAATC	7900
AGCCATGGTG	AGAATATTTA	CCATGGAAAT	TGGCAAATAC	TAAAAAGCAG	7950
AGCACCTTTT	TTTCTGAGAG	CCAGACCATA	GCTCTTCTAC	TCCATAGCAC	8000
CCATCATAAC	AATTTTAA	TACCTCCACT	GAACAGCTTC	TTCTCTCTCT	8050
TACTTCTTCC	ATATCTGATT	TGAGCTTCTT	AATTTATCAT	GTGAACCACT	8100
CTTGTAATAA	TAACCCCAAA	TCCCTGTTC	ATTGTTCTTC	CTGTAAAAAT	8150
ACTAAACCTG	GTTTAGTCCA	ACCATATTTT	CTCTCTTTGG	AATCTACAGG	8200
GTGGCCCAAA	AACCTGGAAA	TGGAAAAATA	TTACTTATTA	ATTTAATGT	8250
ATATTAATAA	GCCATTTTAA	TGCTTCAATT	CCAGTCTCAG	TGGCCACCCT	8300
GTATAGCTGG	GCTATTGAGC	TCTTGGGGGA	GGAGGGAGTG	GACAGTCTCC	8350
CAGCCACACA	GACTGATGTT	GCACCAAAAC	TTTTTTAGCT	TCCAGACTTC	8400
CCTGGCCCTT	AGTGTTACCC	TTAACTCTCC	ATTTCTCTGC	CTTTCACATT	8450
CTCTACTTTT	TAAAAATCTC	TGACTCCACC	TTACCTTAT	CATTCTTAGC	8500
ACATGACCAT	ACTTCTGCTT	CCCAAAGAAA	ATGAGCAATT	ACTTCTTTT	8550
CCTTTCTCTC	CTGTCACTAA	ATCTGCAGAC	ATGTCATGCC	TAAGTCCAGC	8600
TTTCTCTCTT	TCTCTGATCT	CAGTCTGCTT	CTTCCATTTC	TGCCCTGAAT	8650
CCCGTCCCTT	CCCCAACCCC	CAAGGACTTC	GCTCTATCAG	TCACCTCTTC	8700
CCTCTCCTGT	ATCTTCAACT	CCTCCCATTT	TACTGGCTTC	TTCTCAAGC	8750
CTTTCCCCAA	GCCTTTCCCA	TCTCAATTAC	CTCCTCGCAC	ATGCCTCTGC	8800
AGAAACCACC	CCGTTTCTTC	CCTCCCCCTG	GCAGCCTGTT	CTTCTGTTC	8850
TGCCCTCATG	ATGGCCACTC	CATTGTGTCA	CTAAAAATCA	TCTCTCCGAC	8900
ATCATCAATG	GCCTTCTTTT	GTTGGGAAAC	CTAATAAACA	CTTTATCTTA	8950
TTTGGTCTTT	GTTATGGGTT	GAATGAGGTT	ACCCCGAAAT	CCATATTAGA	9000
AGTCCTAACC	CCCAGTACCT	CAGAATGTGA	CTTTATTTGG	GAATAGGGTC	9050
ATTGCAGACG	TTATTAGTTA	GGATGAGGTC	ATACTGGAAT	GTGATGGGCT	9100
GCTTATCTAA	TATGACTGAT	GTCCTTATAA	CAAGGAGAAA	TTTGGAGACA	9150
GACACGCACA	TAGGGAGAAT	ACCATGTGAT	GACAGGAGTT	ATGGAGTTGG	9200
AGTCAAAAAG	CTATGGGAAC	TTAGGAGAAA	GACCTGGAAC	AAATCCTTTC	9250
CTGCGCCTAG	AGAGGGAGTA	TGGCCCTGCC	ACTACCTTGA	ATTCAACGTT	9300
TGGGCTTTTC	AAAACGTGAA	GACAATACAT	TTCTGTTGTT	CAAACCAATT	9350
AGTTTGACAGT	ACTCTGCGAC	TGCAGCCCTA	ACAAACTAAT	ACAGTCTCTT	9400
GGAGGCATTT	GGCAAGGTTG	ACAATGGAAG	CACCTTCTTA	CCCCCTTAGG	9450
TCTGTGCGCT	TTCTTGTGTT	GGGGTGTGTT	CTAACAATT	CTCTCCATCT	9500
CTCTCTCTCT	AGTTTGTCTT	AAACATTGGT	GTTCTTCAGA	CTTCTGACCT	9550
AGGCCTTCTT	TTCACTTCTC	ATATTCCCTT	GGGTGGTCTC	ACCCACTTCC	9600
AGAAATTACT	TAAATTACTG	CTCATGCAGT	ACTGTGCTGG	AAACTGTTTA	9650
ACAACCTGGT	CTCTGGGAAG	AGGGGAGACT	GGTTGATGGT	TTTGTGCTGAT	9700
TTCTGTGGTG	TAAATACTCC	CTCCATGGCC	AATTCCAAAC	TGCCAACAGT	9750
TTAACAACCTG	GCTCACAAAT	TTTCTCCAAA	TTTAACATTT	GGCTTTCACA	9800
GGCCAACAAC	GTGGTACAGC	CAACTCCAGC	ACACCTCTGC	TTTGTGTGTA	9850
GAGAGAAATA	ACTTATTTTT	GTACAAAAGG	TAAAAATAAA	ACACCTGCAG	9900
GCCCCCTTTT	TTTCTTAAAC	AAACTGCTCT	AGAAATAGAA	TAGCTGAAGC	9950
TTCTTTTATG	CATTCTCTCT	TTATTTCCAT	GTCAGTGTGG	TGGTGGGATT	10000
ATTTTCTCTT	TATTTTCTCT	GTATATGGTT	GAAATACTGT	ACCTTTGATC	10050
AGTTTATAGT	TTATGGCATG	TTTTGCACCC	ATATTAAATC	TAGTTTGTGT	10100
CAGAGGGCGT	CAATATTATT	TTCTCAAAAC	AAGAAATAT	TTCATTGCAA	10150
AGGAGACAAA	CAAAAAGGTC	CTTAATACCA	AAACTTTGAA	ATGTGATTTC	10200
TTGTACTTGG	CAGTGTCCAA	GTGGTAAACC	CAAACAGTAT	TGGGTTTTCA	10250
TTTTGTTCAG	GAAAGTCTTT	GTCTGGCAGC	GACTTACCCT	TACATCAGGC	10300
GGGCTTGTCT	CATTCTTCTA	CTTAAGTATT	TATTAACAC	CAGCGGTGTG	10350
CCAAGTACTT	ATCTAGGTAT	CGGGTAGATT	CTGATAAGTC	AGTCAGGTCC	10400
CTGCTCTCAG	GGAGCTTGCA	GCAGAGATGG	GGGCTGCAAT	AGAGAGTAAG	10450
CCAAGGAAAT	GAAAAAGGAA	GTTGATTTCA	GAGAGTGATG	AATGCTATGA	10500
AGAAAATGAA	GGCAGCGCAG	TGTGATGGAG	AGTGACCCAA	GGTGGTACAG	10550
TTTGTACCTC	TAAGGACCAG	ACTGTGACCC	AGGTCACTCA	CAGATGCCCT	10600
TCATGTGATG	CCACAGCAAC	TTTTCCAGGT	GCTCGTTTCC	TCCCACTTCC	10650
CAGTCTCTTG	CCCAGCCGCG	ACTGCTTACA	AATACAGCTA	GAGGAATCTA	10700
AATGAGGTTT	CTCTATCATC	AAACCCAATC	AAAATGCCAA	GGAACAGAAT	10750

CAGTGCCTGG	CTGAAGGCAG	TGGAACAGGG	CCAGCCTGGA	GTGGTTCTCT	10800
CTGAGGAAGT	TCCTCATCTT	GGTTTTAGGG	CCATACCTTG	TGACCTGTGA	10850
GCTAGGGGTT	GCCAGTCCCT	GACATTTCTA	CTGAGGACTC	GCCTGTCTAT	10900
ATTCCCCGCC	TGTATGTGTC	TCCTGAGTTC	CAGACACACA	GGGCGAAGCG	10950
CCTGATGGAT	GGAAGTATGT	TTTTTGGTGT	TCCATTGGTA	TCTCAAATTC	11000
TACAAAACTT	AGTGCCCTTT	CTCCTCCCTG	TTCCCTCCCA	TCTTCAGTCT	11050
ATCACCTGTT	CCTCATCCAG	CAAAATGATAT	TACCATCTTC	CAAGGAGCTT	11100
CCCAGGAGTA	ATCCTTGACT	CCTCCTCAAC	ATCCAATTAA	TAATCAAATC	11150
TAGGCCAGGT	ACAATAGCTC	ACGCCTATAA	TCCCAGCACT	TGGGGAGGCT	11200
GAGGCAGGTG	GATCATTTGA	GGCCAGGAGT	TCAAGACCAG	CCTGGCCAAC	11250
AAGGTGAAAC	CTGTCTCATT	TAAAAAAGT	TATTTTAAAA	ACTCAAATCT	11300
ATTATTTCTA	CCTCTAAGTG	TGCTTGAAAT	TTATCCATCT	CTCTCCATCT	11350
CTGAGCTGTT	ACCTTACCTC	AGTCCATCAC	GTTTTGTCTA	CGTTAACATG	11400
ACCAGAGTCT	TGTTCTTAGT	CTGGTGAGGT	CACTCCAGCT	GCTTCAGATC	11450
CTTCCATGGC	TCACCGTTGC	CCTCATATAA	AGTTGGCACT	CCTGGACATG	11500
TGGCTTACGG	GGCCCTCCGT	GATGTGGCCC	TATTTGCTTC	TCCATTCTGT	11550
TCTCTCCAG	CCTCTCTGCC	CCCATCTCTA	GGCACCACCC	ACACCTTTCT	11600
GCTCGTCAAT	GGTGCCAGCT	TCTCTTCTAT	CTCTGGTCTT	TGGACAGACT	11650
TTTCCCTTCA	CCTGGAATGC	TTTCTTCAAT	CCTACCCAC	TCTCTTTAAT	11700
CTAGATAAGG	TTTATTTCTT	TTGAATGTCT	AGCAGTGAAA	CCATTTCCCC	11750
TGAAAAACCT	TCTCTAACCA	ACCCCTTACC	CTCAGCCCAA	GGTCTAGATT	11800
AGGAGTCCCT	CTGAATGTTT	CCATAGCATT	TTTAAAGAAT	TGCCTATTTA	11850
CTTGTTTCGTA	TCTATCACTA	AACATCAAAAT	TGTATGAGAA	CAGCCACTAT	11900
CTCTGCCCTGG	TCTACCATTC	ATCTCCAGCA	ACTAGCATAA	TGCCTGGCAG	11950
AGTCAGCCTG	CAACAAATAT	TTGTTGAATA	AATTAACAGA	TGGCTTTATC	12000
TCCTTAAGTA	AATCTTGCTT	TTTTACACCTA	TTAAAACAGA	CGCACAGGCG	12050
AGGTGTGGTG	GCCCATGCCCT	GTAATCCAG	CACCTTTGGCA	GGCTGAGGTG	12100
GGCGGATCAC	CTGAGGTCAG	GAGTTCAAGA	CCAGCCTGGC	CAACATGGTG	12150
AAACCCCATC	TCTAATAAAA	ATACAAAAAT	TAGCTGGGCA	TGGTGGTGGG	12200
TGCGTATAGT	CCAGCTACT	AGGGAGGCTG	AGGCAAGAGA	ATCGCTTGAA	12250
CCCAGGAGGC	AGAGGTGGCA	GTGAGCCGAG	ATCATGCCAC	TGTACTCCAG	12300
CCTGGATGAC	AGAGACCCCTG	TCTCAAAAACA	CACACACACA	CACACACACA	12350
CACACACACA	CACACACACA	CACACACACC	AAGTTGTATA	ATTTAAAAATA	12400
TAACGTGCTT	GTTATGGAAC	ACTTGTAATA	TACAGGAAAG	TAATGAAAAA	12450
GTCTACCATC	TAGCTCACCA	CATAATGACC	ATTGCTATCA	TCCTGGCATA	12500
ATTCTCTCCT	GTATATAAAT	ATATATTCTT	TTATTGTATA	AATTACACTA	12550
TGAGTACTAT	TTATTTATTT	TACTGTGGCA	AAATGCGCAA	AACATAAAAT	12600
CTTGCCATTT	TAAGGTATGC	AGTTTGGTGC	ATTACCCACA	CTCACATTGT	12650
TGTGCAATAA	TCACCATAT	CTATCTCAGA	ACTTCTTCGT	CTTCCCAAAC	12700
TGAAACTCTG	TACCCATTAA	ACAATAGTGC	ATCCTCTGTT	TTCCCTCCCC	12750
TACAAATTTAT	TTTTATTTGG	GTTTGTACCA	AACAGAAAAT	AGCTGCTTTT	12800
TCCTTACTTA	GTTTCAGATTA	GCATTTCCAT	TTATTTAGCC	TGTTGTTTTGA	12850
GGATGCCATG	ACAGATGCCA	TCCTTCTCTAG	AGCTCTTTGG	GGCTGTCCAGG	12900
TATTTTCAGT	AGGTGGAATT	CGGGTTGATA	ACATTTTAAA	ATCTCACTTT	12950
ATTCTGAGGT	TCCTAGTGTG	AGAGCCCAAC	GTATTTTATG	GGACTCCCAA	13000
GTTACAAACA	AAAATATGGT	GAGGAGGAAT	CACCTGAAGT	TTAACACAAG	13050
AGACTTACAT	TTTGTTCAT	TTCTATCTTT	TAGTTTATTT	CCTAAGCATA	13100
AAGAAATACT	TTGAAAATTT	TACATAGCAT	TATACATATT	TAATTAAGCA	13150
TGAGCACATC	TTAAAACTTT	AAATTTTAGA	TCAGATCTTT	AATTCCTAGG	13200
ATATTAAAGAG	GTACTGGCAA	TTTGGCCAGG	TGTGGTGGTT	CACGCCATATA	13250
ATCCCAACAC	TTTGGGAGGG	TGAAGTGGGC	GAATTGCTAG	AGCCCAGGAG	13300
GTGGAGGCTG	CAATGGCCTG	AGATCACGCC	ATCGTACTCC	AGCCTGGATG	13350
ATGAGAATGA	AATCCTGTCT	CAAAAAAATA	AAAAAATAAA	AAAAGAAGAA	13400
GAAGAAGTAT	TGGCAATCAG	TGCTCCAGGA	ATAATTCTCT	GACTTGAAAT	13450
AAACCTACAT	GTAGACAAAC	TAATTAGGCC	ATTCCAAGAG	TTGCTAGCAT	13500
TGGTTTAATA	TGTTTTCAGA	GCATTCCAGG	AAGCAGTGTG	GCCAGCATTG	13550
CATGTTTGAT	ACTTCAGAAA	TGTATGACAG	GTGTTTCTCT	TACCCAGGTC	13600
TTCTGTTTTT	TTAGTTTTGC	TCATGTAAAT	ATTTATGAAC	ATCCTCATCT	13650
TTTTGAGGGA	AGGGATTATA	GATCATTTCTA	ATTCCATTTT	CTAGCATTTG	13700
GTACCATTTCT	AAGCACATGA	TAGGCACCCA	TTTGGAGCAT	TTTTGGCTTG	13750
ACAGAATATG	CATTTAGAAT	TGTTCAAATT	AGAGGTGTCA	GTGATGGGAA	13800
TTAGAATACT	ATATAATTCT	AAGTCATTTG	ACTTAAATAC	AAAAGAATGA	13850
TTTTCTTTGG	TGGGGAATGG	TGAAGGGAGG	CAGGAGTTAA	GAAGAGGAGA	13900
AGAGATCCTA	AGTCATTTAT	AAACTTCTCT	GGAAAGACAG	GTGTGTGAAG	13950
ACTTTTTTAAA	AAGTCATPCA	CCAAATTGTG	TGTGTGTGTG	TGTGTGTGTT	14000
TTAAATAGAC	TTTATTTTTT	AGAGCAGTTT	TAGGTTTACA	GCAAAATTGA	14050
ATGCAAGGAC	AGAGATTTCC	CATAAACCCC	CTGCCACAC	ACATGCATAG	14100
CCTCCCTCAT	TATCAACATC	CCCACCAGAG	AGGTGTTTGT	TCTAGTTGAT	14150
GAACCTACAC	TGACACATCA	TTATCACCCA	AAGTCCATAG	TTCACGGCAG	14200
GGTTCACTGT	CGGTGTACAT	TCTATGGGTT	TGAGCAAATG	TATAATGACA	14250
TGTATCCACC	ATTATAGTAA	CATACAGAGT	ATTTTCAGTG	CCCTGCAAT	14300
CCCCTGTTCT	CCACCTATCT	ATCCCTCCCT	CTCTGCATTT	CCACCCACAG	14350
CCCCTGGTAA	CCGCTGATCT	TTTTACTGTC	CCATAGTTTC	GGACGATCTA	14400
TTTTTTCAGAG	AGACACAGAG	CTGTCTTTCC	CTTAGTTTCT	ATTCTATCAT	14450
TTCTTTCTCC	CCATCCATCA	TAAAAGGCTA	TGAGTTTTTT	TTAAGTGTG	14500
AACACCATCC	TACTTGTCAC	GTTAAACAT	AAGCTCCTGG	CTGGGTACAG	14550
TGGCTCATGC	CTGTAATCTC	AGCATTTTGG	GAGGCTGTGG	CAGAAGCATC	14600
ACTTGAAGCC	AGAAGTTTGA	GACCAGCCTG	GGCAACATAG	CAAGACCCCA	14650
TCCCTCCACA	CACAAACACA	CACACACACA	CACACACACA	CACACACACA	14700
CACACACACA	CACAAAAACA	AGCTCTTGCC	AGAATTAGAG	CTACAAATTG	14750

CCCTCAGGTT	CCTAGAAGAT	CAGTCCTTCA	ATTAGATTCA	GATTGAGATG	14800
CTTCCTCTTT	TAAACAATGA	TTCCCTTTCT	ATCATGCCCA	ATAAGAAAAC	14850
AAATAAAAAA	TAAACAATAC	TGCCGTGAAT	CTCAGCTACC	CAGGAGGCAG	14900
AAGCAGAACT	GCTTCAACCC	GGCAAGCAGA	AGTTGCAGTG	AAGTGAGATC	14950
GCGCCACTGC	ACTCCAGCCT	GGGAAACAGA	GCAAGATTCT	GTCTCAAAAA	15000
CAAAACAATG	TGATTTCCCT	CTCTAAGTCC	TGCACAGGGA	AATGTTAAGA	15050
AATAGGTCCA	CCAGGAAAGA	AGGAAGTAAG	AATGTTTGAC	TAGATTGTCT	15100
TGAAAAAAAT	AGTTATACTT	TCTTGCTTGT	CTTCCTAACA	GTTCTCCAAA	15150
GCTTCGTACC	TTGGCCAGAG	GCTTGCTCTC	TGCGTACCTG	AGGTTTGGTG	15200
GCACCAAGAC	AGACTTCCTA	ATTTTCGATC	CCAAGAAGGA	ATCAACCTTT	15250
GAAGAGAGAA	GTTACTGGCA	ATCTCAAGTC	AACCAGGGTG	AAAATTTTTA	15300
AAGATTCACT	CTATATTTTA	ATTAACGTCA	GTCCGTGTCG	AGAATGCTTT	15350
GAGAAAACTG	TTATTTCTCA	CACCTAACAA	TTAATGAGAT	TAACCTCCTC	15400
TCCCCTCATC	TGACCTGTGG	AGGAATCTGA	ACAAGAGGAG	GAGGCAGTGG	15450
GCAGGTTTTC	TTATCATGAT	GTTTGTGTCG	TTCAAGTGTA	GGCCTCACAA	15500
AAAAAATAAA	AAAAAATAAA	GGCGTCCTGG	ATATAACTGA	GAGCTCATTG	15550
TACAGTAAAT	ATTAATAAAA	CAGTGATTGT	AGCTGAAGGA	TAGAAGTCTG	15600
TGGAGGGAGC	AAGTGGGTAG	AATCGCGTCA	AACTAAAGAG	CATTTCTAGC	15650
CAAAGACACA	ATGATAGATT	GAAGGATATT	TATTTCTAAAT	ATAGAATATG	15700
GGTGAACGAG	ATCTGTGGAC	TTCTGGGCTC	CAACGTTAGA	TTCTGATTTT	15750
AGCAAGCTTG	TCAGGGGATT	CTGATATTGA	AAGGCTGTGG	CCTTCACCTG	15800
AGAAACCTGC	CCTAGGGGGC	CATGAAAAAT	TGTCCTGTCT	TTCAGAAGTG	15850
CTATCAGACA	TCAAATGGAA	GTTAAATCGT	ATCTTAACAA	TTACTAGGAT	15900
GGGCGCAGTG	ACTCACACCT	GTAATCCCAA	CACTTTGGGA	GGCTGAGGCA	15950
GGAGGATCAC	TTGAGCCGAG	GAGTTCGGGA	CCAGCCTGGG	CAACATAGAG	16000
AGACGTTGTC	TCTATTTTTT	AATAATTTAA	AGAGAAAAAA	ATACTGAAAA	16050
TATTTGTATC	ACCACCTGAAT	TATAATAATG	TGTATATAAT	GTATATATTCT	16100
ATTATGAGGA	ATATTTGATT	ATTTTCATATA	TTATATCTTT	TCCTTCTGTT	16150
TATTTTTATCC	AGTTATGAAG	TATTTAGAAC	AATTCATCAG	TAATTTGGGGC	16200
TAAATTGACA	GAATAGTAAT	CAGAGAAAAA	AGAAAAAGAC	AGATGGGTTA	16250
TCTTTGAATA	CCAGGTTGGA	GTTGTTTATG	GGTTTGTGTT	TTGTTTGGGG	16300
GGCGTTTTTT	TAGACAGAGT	CCCCTCTGTC	TGCCCCAGGCT	GGAGTGCAGT	16350
GGCACAAGCA	TGGCCCACTG	CATCCTTGAC	CTCTTGGGCT	CAAGCAATCT	16400
TCCCACCTTA	GCCTCCTGAG	TAGCTGGGAC	CACAGGTGCA	TGTCACCACA	16450
CCCAGCTAAT	TTTTTTATTT	TTTGTAGAGA	CAGTCTTTCT	ATGTTATCCA	16500
GGCTGATCTC	AAACTCCTGC	ACTCAAGTGA	TCCCCCTGCC	TTGGCGTCCC	16550
AAAGTATTGG	GATTATAGGC	ATAGCCACCA	CACCCAACCT	AGTTTCTATT	16600
TAGACTTGGC	CCTTTCCCAC	CAGTCATTTG	TGTCCAAAAG	ATCTCATAAA	16650
TGTAGACAGG	AAACTGTCCCT	TTGCTCATCA	GTTTTCTTCA	TCCTGTGTCT	16700
AGGGGGATGG	TCGGTGGGGG	AAACTGGGGT	TATGCAAGTT	CCTCTGAAAC	16750
ATCCTCTGTG	AGCCCAGGGA	TGGATGAGGC	ACCAGCCGCC	AGCGAGTCAG	16800
TGTGCAGCTT	TCCAGAAAGG	AAGTCATCAG	CCAGTCAGCC	GGCCCTGGCA	16850
GCCAGCACCC	GGCAACCCCT	CTGTCTTGTC	ATAAAGAAAT	GGTCTGCCTG	16900
ACAGGATGGT	GTGGATTTTT	CTTTTTTCTT	TTTTTTTTTT	TTGAGACAGG	16950
GCTGGCTCTC	GTGCCCCAGG	CTGGAGTGCA	ATGGCGGGAT	CCTGGCTCAC	17000
TGCAGCCTCT	GCCTCCCAGG	CTCAAGGCAT	CCTCCCACCT	CGGTCTCCCG	17050
AGTAGCTGGG	ACCACAGGCA	CACACCACCA	CGCCCAACTA	AGTTTTCGTA	17100
TTTTTAGTAG	AGGCAGGGTT	TTACTATGTT	GTCCAGGCTA	GTCTCAAACT	17150
CCTGAGCTCA	AGCTATCCAT	CTGCCTTGGC	CTCCCAAAGA	GCTGGAATTA	17200
CAAGCGTGAG	CCACTGTGCC	TGACCAGGGT	GGATTTTTTC	AAGTGCACAT	17250
GTTGTGGTCC	CAGAAGCTCT	GATGGTACCA	AATTCCAAGC	GAAAAAAAGT	17300
CAATGGTTCC	CACCCATCCT	ACCTCCCCTG	ATGGCAAGAG	GAAATCACCA	17350
CACCTGAGAT	ACAGTCCATG	TAAACAAAT	TGCTATGGAT	TTTGAAAGTG	17400
AACCTTAAGA	GAACCTGCAT	ATGTTTTCTT	CATTAGAGTT	CTCTGGTAAT	17450
TTCCAGCTTT	TTTTTTTTTT	TTTTTTAGAC	AGTGTCTCGC	TTTGTGCCCC	17500
AGTGTCAACC	AGGCTGGAGT	GCAGTGACGT	GATCTCGGCT	CACCTGCAACC	17550
TCCGCCTCGT	GGGTTGAAGT	GATTTCTCTG	CCTCAGCCTC	CTGAGTAGTG	17600
GTATTTTAGT	AGAGACGAGG	TTTCACCAT	TGGCCAGGCT	GGTCTCGAAC	17650
TCCTGACCTC	AAGTGATTCC	CCCATCTCAG	CCTCCCAAAG	TGCTGGGATT	17700
ACAGGTGTGA	GCCACTGCAC	CCGGCCAGTA	ATTTCAAGCT	TCTGAGGAGC	17750
CCTTTGAATT	GTTAAATAAC	TTGTAGCTAT	GTCCAACATA	TCCATGTTCA	17800
GTGTATGTTC	GATATTTCTT	AGGAAACCTG	CCCTTGGTTG	TTTTCTTTGT	17850
GGTAATTCAT	GAGCCGGCAA	ATTTGACATG	TGTTACAGAA	TATACCTTTT	17900
CTCTGCTCTC	CTACCTCATA	ACCAGAACTT	AATTATCCTG	CTTTAGTCAC	17950
ATAAATAGCT	AACTAAATAA	ATATATGAGA	TTTCAGTCTG	CTCACTGTGA	18000
AAATAGACCT	TCTAAATGAT	CTCTTCCACT	TGCAGATATT	TGCAAAATATG	18050
GATCCATCCC	TCCTGATGTG	GAGGAGAAGT	TACGGTTGGA	ATGGCCCTAC	18100
CAGGAGCAAT	TGCTACTCCG	AGAACACTAC	CAGAAAAAGT	TCAAGAACAG	18150
CACCTACTCA	AGTAAGAAAT	GAAAGGCACC	CTAGAGATGT	TCCAGCCCCA	18200
AAGATATTTG	AATAGGTTGG	ACTCGGGCAC	CAATCTAGCA	AGTCTACGCG	18250
AAGTTGTATA	AAGCTGAAAA	TACTGAAGCA	TTTCCCAAAAT	GGGAAATCCT	18300
AAACTCAAAA	CTTGCTTTTT	GGTTTTTTTG	TTTGTTTGTT	TTTTTCTCAT	18350
CTGACATTGC	TTAGTAGTCA	CAGAATGAAA	GATAAATCAA	TCATTTCATGA	18400
TCTAACAAATG	ACCTTCAGTG	CTCTAAAAAA	CTACGGAGTC	AAGGAAAAACA	18450
TGAATATATT	CCTCATGTAA	AATTAATAAA	CAGACATATA	AAGGGCAAAA	18500
CATGAACATC	ATTCTAACCT	TGAGGTCCGT	CCCCCTCCCA	GAAATAACCC	18550
CCAGTATGCC	TTGGTTTAGA	GCATTAAGCA	GGAGGGCCCT	GAGTCACTCC	18600
AGACAGTCTT	GACCACCAAG	CAGCATTCTC	TTTTTGTGTT	CTCTGTGGTG	18650
TTTGCAAAACA	CAGGGCTAGC	TCAGCTACCC	ATTAGTATGT	TTTCAGTCAC	18700
TAAAAACAGTC	TTCCAGTCTT	CAAATTAGGA	TGACATTGTC	ACATGGGGCT	18750

TTAAAGCAAG	TGAAACAAGG	AACCCCTTT	TTTTTTTTTT	TTGAGATGGA	18800
ATCTCACTCT	TGTCGCCCCAG	CCTGGAGTGC	AATGGCGCAA	TCTTGGCTCA	18850
CTGCAACCTC	CACCTCCCAG	GTTCAAGAGA	TTCTCCTGCC	TTAGCCTCCT	18900
ATTCAATTATG	AGGAATATTT	GATTATTTCAG	TTCTGTAGG	GTAAAGATAT	18950
TACCCCCGAT	CATATTATTG	ATTATTGAGT	AGCTGAGATT	ACAGGTGCCT	19000
GCCACCACGA	CCGGCTAATT	TTTTGTATT	TTAGTAGAG	ACAGGGTTCT	19050
ACCATGTTGG	CCAGGCTCCA	GGCTCGTCTC	GAACCTCTGA	CCTCAGGTGA	19100
TCCACCACC	TCAGCTCCC	AAAGTTCTGG	GATTACAGGC	GTGAGCCACC	19150
ACTCTGGCC	ACAATCCTTT	TTTAACTATG	AAATATATTT	TTATCTGAAG	19200
TTTGATGTTT	ATACCCAAC	GAGGGATGAT	GTTCCTATAT	CTCAGTTAAA	19250
GAAATAACCT	GCTCAGATAC	TTCAAGCTCT	TCTTTTGACT	TTTGAAAATA	19300
AATGATCTTG	AAGTTACTAT	ACTTTGTTTG	GGTAGTTAA	CATTATTTAA	19350
AGTATATTAT	TTTAATTAAT	TATCTTTGTA	AGATTTTACT	GTATACTACC	19400
TGGAGTTCAA	TGTATCAGAT	GGATTTCAAA	TTTATGTACA	TTTTTTATGT	19450
ATATGGTACA	GAAAAAATG	TGATCCATAA	GAAATCAGAA	AATAGCGCAT	19500
ATGCTAATAG	CTAATGTGTG	CCTCTAAAAA	ACTTATTTTT	GCATTTTTAA	19550
GAGGGGGATA	TACTCTGACA	CTTTAATAAG	TGTAATTAAT	TATTGACTGG	19600
AATTTGGCAT	GAGGCAGGGC	CATTTCAGAT	CCCATTAAAG	GAATGACACA	19650
TACCAGAGAA	CCACAGAAAG	AAGGCCACAT	TTGTAATAAA	TCATTATAGC	19700
TCTGCTAGGA	GAAAGCCAG	TTGTATTAGG	TAATTAATGG	ATTTGCTCTT	19750
AAAACACATG	TCCCAGGAAG	TATAGGTGAG	TCTTGGGGGG	CCGCATTAAA	19800
CATTATACCA	ATGTATCTTA	CATTCTAAG	AAAGTTTAC	TACTTTACAG	19850
GATCTTTCTG	TTACCAAAAT	GGAAGGTTTC	CAACTCCAGG	ACTTGGCTTT	19900
CATAGTTCTT	ACACCAGGGG	AAATGCCTTC	CTTTGCTAAC	TATGCAACCA	19950
GGTTAGTTAG	TGTAAGTCCA	GCCACCCTGT	TGGCAATGCT	AAAAGGTACA	20000
ACAAACACAG	AATTTTATTT	GCATTTGTAA	ACATTTGATT	TCTGGCTCGA	20050
AATTTTCAGT	TTTCATGGGG	ACGTCATGGA	AACAGAAATC	TTCTGTGTTT	20100
AGTTTGGGCA	CCTACTCAT	GTAGTGACAA	ATATTTTCAGA	AGCCAATAGG	20150
GGATTCCACA	AATTTGTCTG	AACCTGTGGC	TGAGACTGGT	AATGGCTGAG	20200
TGACATGGGG	ACATACCACA	AAAGAAGAGG	TAGCAAAAGG	CTGCTGAGAT	20250
AAGGACATGT	TCATTGCTTA	GCTAGTGGCC	TGCACCCTTA	AAACACATGT	20300
CCCAGGCTGG	GTGCTGTGGC	TCACGCTGT	AATCCCAGCA	CTTTGGGAGG	20350
CTGAGGCGGG	TGGATTACCT	GAGGTCAGGA	GTTCGAGACC	AACCTGGCCA	20400
ACATAGTGAA	ACCTCATTTT	TACTAAAAAT	ACAAAAATTA	GCCAGGCATG	20450
GTGGCGGGCG	CCTGTAGTCC	CAGCTACTCA	GGAGGCAGGC	AGGAGAATTA	20500
CTTGAATCTG	GGAGGCAGAG	GTTGTGGTGA	GCCGAGATTG	CGCCACCGCA	20550
CGCTAGCCTG	GGCGACAAAG	TGAGACTCTG	TCTCAAAAAA	ACAAAAACAA	20600
AAAACAAACA	AACAAAAAAC	AACAACAACA	AAAAAACGGG	TATCCAGAA	20650
GATACAGGTA	AGTTTTCTAA	CACAGGTCTT	CTTGTATGGT	GCGTTCCACT	20700
TAAGTAGAAG	ATGACAAAAA	CATTGTGCTA	GAGAATATAG	ACTCACATTT	20750
TAAACCTGTT	TGAGCAGGAA	AAGGAAGCAA	TGTTACAGAT	GTAAATCTGG	20800
GTGTGACTGC	AGAAAGGATG	ACTCCCTTAT	TAAAGTAGTC	ATCCTGAGTG	20850
AGCTAACTCT	TTGTACTTCC	TCTTCTCCTC	CTGTTCCCTT	CATCACCCCA	20900
TTCTTCCGTT	GCCTACACCC	AGGCCACAT	TGGATGCTGA	CATAGACTTA	20950
CATGGTACAG	TCCAAGGGAA	AGATCTGCCA	TTTTTTTCAA	TGTGTCATCT	21000
TGGTTATCTT	CATTCCAAAG	ATCTCTCCAC	TCTTTATACA	GTAAGAGATG	21050
AGAGTCTGGA	AAGGATTGGG	AATAAGATAA	TGAATTGTAA	GTTTTAAATT	21100
GTTCTTCGTA	TTTTGGGGAA	GGAGTAGGCT	AGGTGGTCTT	TCTGTTTTTT	21150
TTTTGTTTTT	TTTTTTAAAG	TAGATGTGGC	CAGACGTGGT	GGCTCACGCC	21200
TGTAATCCCA	GCACTTTGAG	AGGCTGAGGC	AGGTGGATCA	CTTGATGTCA	21250
GGAGTTCAAG	ACCAGCTTGG	CCAACACAGT	GAAACCCCGT	CTTTACTAAA	21300
AATACAAAAA	CTAGCCGGGC	TTGGTGGCGT	CCACCTGTAG	TCCCAGCTAC	21350
TGCAGAGGTG	GAGGCAGGAG	AATCACTTGA	ACCCGGGAGG	TGGAGGTTGC	21400
AGTGAGCCAA	GATCATGCCA	TTGTACTCCA	GCCTGGGCGA	CAGAACAATA	21450
CTCTGTCTCA	AAAAAAGAAA	GAAAAAAGAA	ATGGATTGTA	21500	
ACTCAGTCGT	CAATAGCCTC	TATTCCAGGA	GATGTTACAG	TTGATTATGT	21550
TATAGGGGGT	GTATAATAGA	ATTTTCAGCT	ATGTAATTC	CAAGTGCATT	21600
TGGAAGAAATG	AAGAAATGGA	GGAAGGGTAA	AGTATGAGTG	CAAGCATTCC	21650
AGGTTTTTTG	AAAATGCTAT	AATCTTTGTT	CAGGGCTAGT	ACAAAGTGCT	21700
ATTTAGCTGT	AAGGGTTTTT	TGTGATTTAC	AGACAGTTTT	CACATGTGTC	21750
ATTTCAACCT	TGGTTTTATG	GCGAAGGCAT	GTGATGGTGC	TTGTCCCAGG	21800
ACTTTAGATC	CATATCTGAG	GTTCTGTGCG	GGCAAAGATA	TTACCCCTGA	21850
TCATATTATA	GTCTATAAGT	GGGAGAGTTG	TGCCTGGAGC	TCAAGTCTTA	21900
TGATTTCTGA	TCCAGGGCAC	TTCTTACAAC	ATGATTTTGC	AATATAAAAG	21950
CCTATAATGT	GTGACTAAAG	CAGGTCATC	ACCCCTTGTA	ACAGACTCTA	22000
GTAATGGTAC	TGCCACCAA	CGGCTGCGTG	ATATTGGGCA	AAGACTTACC	22050
TTATTTGAAT	TCAGTTTTC	TCCTAGAAAA	ATGAGGGTGG	AGGTAAAGCA	22100
TAGGCTGATG	ATCCTAAAGC	CTCCATACTG	CCCTAAACTG	TGGCTCTAAG	22150
ATCCAGTAGA	ATGCTGGGTC	ACAGGACTCT	AGGGAGCTTT	TCAAACCCAA	22200
ATGTCTGTCA	TTCTTTGATG	GTAGGCAGCA	GTTTATGGAA	GTGGGCGACA	22250
CAGCAAAATAT	CAAAATACCT	AAAGCAGCTT	GCAAGAGTTG	TTTCTGCCTA	22300
GTGGTCTTTA	TAGTTAATAT	TAAATAGTTA	ATTTTTTTTT	TTTTTGAGAC	22350
AGAGTCTTGC	TCTGTTACCC	AGGCTGCAGT	GCAGTGGCAC	AATCTCGGCT	22400
CAGTGAACCC	TCCACCTCCC	GGGTTTGAGC	AATCTGTCT	CAGCCTCCCA	22450
AGTAGCTGGG	ACTACAGGTG	CATGCCACTG	CACCCAGCTA	ATTTTTGTAT	22500
TTTTAGTAGA	GACGGGGTTT	CACCATATTG	GGCAGGCTGG	TCTCGAAGTC	22550
TTGACCTCAG	GTGATCCACC	TGCCTCAGCC	TCCCAAAGTG	CTGGGATTAC	22600
AGGCATGAGC	CACTGCACCC	AGCTTAAATA	GCTAATATTT	AATATTATTC	22650
TATAGTTATT	CAAGTAATTC	AGGCCAAAGA	CTTAGAAACA	AAACAAAAAG	22700
CCACTTTTAA	GGAGAAAGGG	TGTAAGTTTG	CCAGATAGAT	AGAGATCTTT	22750

CTTTTTTAAC	TACAAGAGTT	CAGGAATGAA	TTACTCTTTA	ACAAACGACT	22800
ATAGATATAC	ATGAAAATTG	GAAGGACTTA	TTATGCATAT	GATAATCAAT	22850
TTAAAGACAA	CACCTAAAAT	TATATTGTTG	CCACTCTCAA	AAAGTGGTAA	22900
TAGAACAGCT	AATGGTTTAA	AAAGCAGAGT	ACAGAAGTTC	CCAAACTTAT	22950
GGCACCTTAA	TATCGCAGAA	AACCTTTTAA	AGCATGCCTA	GGCCACAAAA	23000
AATACCTGTA	TTTTGATTAT	TAAATTGTAA	GCTCTACACA	ACCTAATAGT	23050
AATAGGTCCA	ATAGTAATGC	TGTCCAATAG	ATGTTGATGT	TTTTTCTCCT	23100
GCAAACTTAA	AAGATCCCTAC	AGTGCCTCTG	TAAATAGCAC	TGCCTGGTTA	23150
GAGTTGAATT	TCAGATAAAT	AATTTTTC	ATGTTAATTA	TTTTTCTTTT	23200
CTTTACTTTT	TTTTTTGTTT	TTTTGT	TTGTTTTTTT	TTTTGAGACA	23250
GGGTCTCAT	CTGTTGCCCA	GGCTGCTGTG	CAATGGCATG	ATCATGGCTC	23300
ACTGCAGCCT	TGACCTCCCT	GGGCTCAGGT	GATCCTCCCA	CCTCAGCCTC	23350
CCAAGTAGCT	AGCTGGGACT	ACAGGTGCTT	ACCATCATGC	CCGGCTAATT	23400
TTTGTGTTTT	TTGTAGAGAT	GTGGTTTTGC	CATGTTGCC	AGGCTGGTCT	23450
TGAACCTCTG	GGCTCAAGTG	ATCCGCCCGC	CTCGGCCTCC	CAAAGTGCTA	23500
GGATGACAGG	CATGAGCCAC	TGCACCTGGC	CCCTGGGCGA	AGTATTCTT	23550
AATGTTTACA	TAGGACATAC	ACTAAACATT	ATTTATTGTC	TATATGAAGT	23600
TCAAGTTTAA	CTAGGTGCCC	TGCACTTTTA	GTTGCTAAAT	CCTGTAGCTG	23650
TACCATGCA	TTCACTGGTG	CTCCCCAGCT	TGCCTTGCAC	AGAGTTTGA	23700
AACCATAGTC	CTATAACTCT	AGGCCAATT	TTTAATGTAA	AATTTGATTC	23750
ATTTTAAATT	AATAAATAAT	AACAGGAATT	TTTTTAAAA	TTGTTTTAAA	23800
TATAATTAAA	ATTTATCAAAA	TATTTTTTAA	CTGAACCTGT	GACTAGAGAT	23850
ATTTAGATTA	TGAAGAGTGG	GGTTTATGCT	AACTAATGAC	AGTCTGGCTA	23900
TGCATGTGGA	GCACTGAGCT	ATAAATTGTG	GCTTCCCCAA	TTCTCCTGAT	23950
GTCACCTGAA	CAAAACCTAA	GTGTACAGAC	AGAGCTTCTG	GTATCTTCCA	24000
TGGGATTTCA	TTCAACAGCT	GGAGCAAATG	AAGTCAGATT	GATTTTTTTT	24050
AATTTGTCCA	ATTTTGTGTT	CTCAAAAACA	TAATTATAAT	CATTTATTAG	24100
AACTAGAAAT	TCTTCAGTTT	AACAACAGAA	ATAGTTATTCT	ATTATGAAAA	24150
GCGAATCTGG	AGGCCTTCAT	TGTGGTGCCA	ATCTAACCAT	TAAATTGTGA	24200
CGTTTTTCTT	TTAGGAAGCT	CTGTAGATGT	GCTATACACT	TTTGCAAAC	24250
GCTCAGGACT	GGACTTGATC	TTTGGCTTAA	ATGCGTTATT	AAGAACAGCA	24300
GATTTGCAGT	GGAACAGTTC	TAATGCTCAG	TTGCTCCTGG	ACTACTGCTC	24350
TTCCAAGGGG	TATAACATTT	CTTGGGAACT	AGGCAATGGT	GAGTACCCCA	24400
GGGAACAATT	CATTATAAAG	GAGATTCCCC	ACTAGCATT	TTTCTTTTCT	24450
TTTCTTTTTC	TTTTCTTTT	TTTTTTTTTT	GAGACAGAGT	CTCGCACTGC	24500
TGCCCAGGCT	GGAGTGCAGT	GGCGCCACCT	CGGCTCACTT	GAAGCTCTGC	24550
CTCCCAAAAC	GCCATTCTCC	TGCCCTCAGCC	TCCCAGTAG	CTGGGACTAC	24600
AGGCACCCGC	CACCGCGCCC	GGCTAATTTT	TTTTTTTTTT	TTTTTTTTTT	24650
TTTTTTTGCA	TTTTTAGTAG	AGACGGGGTT	TCACCGTGTT	AGCCAGGATG	24700
GTCTTGATCT	CCTGACCTCG	TGATCTGCC	TCCCTCGCCT	CCCAAAGTGC	24750
TGGGATTACA	GGCGTGAGCC	ACCAGGCCCG	GCTAGCATT	TTTCTTATGA	24800
CACTTTTTTT	TTTTTTTTGA	GACGGAGTCT	CGCTCTGTG	CCCAGGCTGG	24850
AGTGCACTGG	CGCATCTCG	GCTCACTGCA	AGCTCCACCT	CCCAGGTTCA	24900
CGCCATTCTC	CTGCCCTCAGC	CTCCCGAGTA	GCTGGGACTA	CACGCACCCG	24950
CCACCACGCC	CGGCTAATTT	TTTTGTATTT	TTAGTAGAGA	CGGGGTTTCA	25000
CCGTGTTAGC	CAGGATGGTC	TCTATATCCT	GACCCCATGA	TCTGCCCGCC	25050
TCCGCTCTCC	AAAGTGGTGG	GATTACAGGC	GTGAGCCACT	GCGCCCGGCC	25100
AACACTCTTT	TTATTTATAG	CAAAATATACT	TCTGCCTGGG	CACATTCTTG	25150
CAAGTGTCAA	ACAAATGCAAC	TTTTGGAGT	GCAATGTGCA	GAAACTCCTG	25200
CTGTATTTAT	TCCAGAACCT	ATTATTGCTA	ATCCCAGTTT	ATGTACATT	25250
TGAAGTGAGA	ACCAGTTGGA	GCCAGCAACG	TTCCCAGCTC	CAAAGTTCCC	25300
TTGAGATTTT	CAGAATCACT	TAACCTATT	ATGCTTGCCA	ACCTGGACTC	25350
AGCAAACTG	GGAAGTCAGC	AGTTTGT	ATTCATCCCT	TCCTTCTCA	25400
GTCTCTCAAA	TGTGTGAGT	AATCTCAGTA	ACCCCATTGC	AACCTTCATT	25450
ACCTGCCCAA	CGGCTTAGA	ACTTGCCAGT	ATAGAATCCT	ACGTGGGTCA	25500
AGCTCCTGAC	TGCTCTCTTC	TTCACCTTT	TTTTGCAAAG	AACTGTAA	25550
TTTTAACTAT	AAGTATTCAT	GATTCGCCAC	ATTTATTCAA	AACATAGAGT	25600
GCTTTTTTCCA	CATATCAGCC	AATGGAAATA	AGGATTAAT	GGGAAATGAA	25650
ATGTAGTAAT	AGGATAAGCA	CAAGTCTTCT	TCTGCTCAA	ACTTTTTTTT	25700
TTTTTTTTTT	CAGACAAGAT	CTTGCTCTGT	TACCCAGGCT	GGAGTGCAGT	25750
GGCGTGTTCA	TAGCTCAATG	TAACCTCCAA	CTCCTGGGCT	CATGCAATCT	25800
CTCACACCTC	AGCCCCCTGA	TTAGCTAGGA	CTACACTATG	CCTAGCCAAT	25850
TTTTTTTCTT	TGTCTGGTT	GTGTGCCCA	GGCTGTCTCG	ATCTCCTGGC	25900
CTCAAGTAAT	CCTCCTGCCT	CGGCCTTCTA	AAGTGTGGG	ATTATAGGCA	25950
TGAGCCACTG	TGCCCGGTCT	CAAACTTTT	TTTCCAAAGT	AAATGAAGTT	26000
ATTAGATATG	GAATATAGTC	TAGTCCCAG	ATATCCATAT	CCATPGGTTT	26050
ATTACCTCTA	TTATTAACCT	CAAATTGTTT	AATAGACCTC	CATATCTCAG	26100
TTATACAGTT	AAAATTTTGT	TTTTGT	CTGGAGTATC	TTATTTATAA	26150
CTATGAGTTT	CACCTTACTT	ATTTATTTTA	TTTTTTGAGA	CAGACGCTTG	26200
CTCTGTCACT	CAGGCTGGAG	TGCGGTTGCG	TGATCATGGC	TCATATGGC	26250
CTCGACCTTC	TGGGCTCAAG	TGATCCTCTC	CCTCAGCCTC	CCAAGCTGAG	26300
ACTACAGGCA	TGCACCAACA	CATCTAGCTA	ATTTTTTTTT	TTCCCATGG	26350
AACAAGGCTT	TACTATGTTA	CCCAGAGTGG	TCTCAAACCTC	CTGGCCTCAG	26400
GGGATCCTCC	TGTCTCAGCC	TACCAAAATG	CTGGGATTAC	AGGCATGAGC	26450
CATAGCGCCA	GACCTGGTTT	TACTTTTCTT	GACTTTGAAT	TACAAGTTT	26500
TGTAATTGGA	AAAATGTTTT	GTTGCTTTTA	AATACTGCTG	TATGTTTGCT	26550
TTTAAATACA	ACATTCTCTG	ATATATATTT	TGAGAATTGC	TGTCTTTCAG	26600
AACCTAACAG	TTTCCCTTAAG	AAGGCTGATA	TTTTCATCAA	TGGGTGCGAG	26650
TTAGGAGAAG	ATTTTATTTCA	ATTGCATAAA	CTCTAAAGAA	AGTCCACCTT	26700
CAAAAATGCA	AAACTCTATG	GTCTGATGT	TGTCAGCCT	CGAAGAAAGA	26750

CGGCTAAGAT	GCTGAAGAGG	TAGGAAC TAG	AGGATGCAGA	ATCACTTTAC	26800
TTTTCTTCTT	TTTCCTTTTG	AGACAGAGTC	TCACCTCGTC	AGCCAGAGCT	26850
GAGTGCAGTG	GTACAATCAT	GGCTCACTGC	AACCTTCGAC	TCCCGAGGCT	26900
AAGCAATCCT	CCCATCTCAG	TCCCACAAT	AGCTGGGACT	ACAGGTGCAC	26950
ATCACCACAC	CTGGCTACTT	TAAAAAAAT	TTTTTG TAG	GATGGGGTCT	27000
CCCTGTGTTG	CCCGAGTCTG	TCTTCTGAAT	TCCTGTGCTC	AGACCATCT	27050
TCACCATCTC	CCGCCCAGAG	TGCCAGGAT	ACAGGCATGA	GCCACCACT	27100
CCAGCCACCA	CTTTTCTTAA	AAAAAAAAT	AGATTCTCTC	TGGTAGACAA	27150
TCCTCAATAG	TCCCATAGTT	ATTTAAACAAT	CTGCTGCCTG	AATACATGAT	27200
TACCAAAAAA	AAGGAAATTT	TGACGGGTT	AGAAATACAA	GGATCTGAG	27250
GCAAAATGTC	CCTATGATAA	AATTGTCTAT	CAAAAATTAGG	AAAGTTGGTG	27300
TATCCTGATC	CTAAAGCAGT	AACCCAGCCA	TTTCTAGGGA	ATAAAACTCT	27350
CATCGGTATA	TTGTGCATAT	ATATGATAT	TATAGCTGAG	TGATAATAAA	27400
ATTTTTTTTC	TAGCTTCTGT	AAGGCTGGTG	GAGAAGTGAT	TGATTCAAGT	27450
ACATGGGCAT	AGTAAGTATG	TCTCCTATG	TTAATACTAG	GAAAGTAGG	27500
CTAGCTTTAT	TTATTACCTA	GTATTCAAAA	AGTTAGTTC	TTTAACTGCC	27550
AATTGACTGC	AGTTCAAATA	AGAAACAAT	AGTGCTCTCA	GTAGCAGCTG	27600
ACTCCAATTT	TAATATTAAT	AAAAAAAT	TTAAGTTAT	TTAAATTAAT	27650
TAGTGGTTTC	TATAAAGAT	ACTTTATACA	GAAAGACAGT	GCCAAATTAAC	27700
CCATGGAACA	TATAAGTAGT	TAAAACCAAT	TGCTTGCCAA	AGAACCAGTA	27750
ACCCAGGAGT	ACATGTCTCT	GCCACTGTGT	TTTTTCAAGA	CAGAGTAAC	27800
GATTTCTAGT	TACTTGCATA	GAATGGAGCT	CTCCTCATAA	TCCTCCTCCA	27850
TCTTGTGCTT	TCCCTAGTAG	AACCTCTACC	TTTTTTTAGT	AACAGGCTGAG	27900
TGGGAGAGGT	AAGAAGGAGA	ATAAGGCTCAG	CAATTAACT	AAAGCAGAA	27950
AGTAAATTT	GTTATTTTTT	TTTCTGAATAT	TTTTCTGTGA	ATTTAGCTAC	28000
TATTTGAATG	GACGGACTGC	TACCAGGGAA	GATTTTCTAA	ACCTGTAGTG	28050
ATTTGACATT	TTTATTTTCT	CTGTGCAAAA	AGTTTCCGAG	GTAATAGTCT	28100
TTTTAAACTT	TTTTAATGTA	AACCAAGATC	CTATTTTTAT	AGTCTAGCTA	28150
GTTCTAAATT	CTATAGGTAT	GTATATTTAC	ATGTTTTTCT	AAATTTAGAG	28200
AACAAGCACT	ATGACTTATC	CACTGTGTAG	TTTTCCCTAT	GCATTGGGTC	28250
TGTCCTCATG	TACGTGATTA	GAAATTTGAA	ATATTTCCAA	TAGCCTTTAG	28300
TAGAAATTAC	TCACATAGAT	GATAAGAATG	AGTTGGTTCA	CTTCACTGTC	28350
CTTCCACAG	CTACTATTTT	CAATAAAGAA	GTTTTCCCAA	GACCTAAATG	28400
ACTATGAACA	TATTTTATTA	TATATAGGA	GGGGTGGGTC	TAGGAATACA	28450
AAGTTTTGAA	TGCTGTTAAT	CTTCAACACC	ACAGTTGAAA	CCACAGGCTA	28500
GCTTTTGTG	AAATACCATT	GATACTTTTT	TGTTCTATAG	TGTGTTGAGA	28550
GCACAGGTCG	TGGCAAGAAG	GTCTGGTTAG	GAGAAACAAG	CTCTGCATAT	28600
GGAGGCGGAG	CGCCTTGCT	AGCCGACACC	TTTTGACAGT	GCTTTATGTG	28650
AGTGAAGCAG	CGCTGGCCTT	AGGGGTCAGA	GTGCACTCT	TCTCCATCCT	28700
TCATTCTCTG	TGAAATAGCT	CCCCAGCCAA	AAAGCAGATC	AAAGACCGTT	28750
TCAGTGGCTG	AGCCCCAAAA	TTCATGCCAG	ATTTTGCAAG	AAAATGATTT	28800
ACTAAAGCTG	GAGGGACATC	TTTTAACAA	GTTTCCAAT	AATCACTATA	28850
AGGATGAATT	GTTTCAGAAA	TTTTGGCCTT	TAATTATGGC	CCATAAATAT	28900
CTCAAGTAGT	CCTTACTCTA	AAGAAGTACA	CTGTAAAAAG	ATGCATATAG	28950
CCGGATATAG	TAGTTCCTGT	TAAATCCAAT	ACTTTGGGAG	GCCAGGTTGG	29000
GAGGATTGCT	TGAGCCCAGG	AGTTTGAGGC	TGCAGTGAT	TATGATGGTG	29050
CCACTGCAT	CTAGACTGGG	GCACAGAGTG	AGACTGTCTT	TTTTTTTCCC	29100
CTCTGTCAAC	CAGACTGGAG	CAAGAGTGGC	CGATCTCAAC	TCACGTCAAC	29150
CTCTGCCTCC	CGGATTGAAG	CGATTCTCCT	GCCTCAGGCT	CCTGAGTAGC	29200
TGGGACTACA	GGAGTATCAC	CGCATGGGGC	TAAATTTTGT	ATTTTTAGTA	29250
GAGACGGGGT	TTTGACATGT	TGCCAGGCT	GGTCTGAAC	CCATGAGCTC	29300
AAGTGATCTG	CCTACCTCAG	CCTTCCAAAA	TGCTGGGATT	ACGGACATGA	29350
GCTACCACGC	CCGGCCACAC	CTGCTCTCTT	AAAAAATAAA	AAATGCAAG	29400
TTAGAGCATA	TTACAGCTTT	GTCTCTCAGG	AGGATACTTA	GTGTATGTAG	29450
CTATAATTCA	TAGATTCCCA	AGAAGTTTAG	AGCCTAAAGT	ATGAGTCCCT	29500
ACCAAGAGGG	CTATCATTTA	ATTTAAAGT	TTGTTAAATC	ATCTCATTGT	29550
CCAAACACAC	AAACTTGATT	GCTTTAAAT	ACTGGTTTAG	TTACATTTAG	29600
TAACCTTAT	AGTGCTTTTA	ATCTATACTG	CTATCTCTC	ACATTGAGAT	29650
TTTTTTTCTT	TTCTCTTCCA	CTTCTATTCT	TTTTTCTCTC	ATCTCATTC	29700
TTATAAGCCT	AGAAATACAT	ACAAATCCTT	TATGCCCATG	GAAAGCAAG	29750
GAATAAGAAA	TGGAGATGTT	TGTTTTGCCA	TTAAGTAAAG	ATCTGGGGTG	29800
TCGGGGAGAA	GGGGGATAGA	GAAAGGAAG	TGGGAAGAGG	TGTCCTAAT	29850
AGCTTAGGTG	CAATTCTGCT	TTTTTTACAT	TTTACCCCG	CTGACTGCCA	29900
CTTTTCTCT	AGCCCTCACA	CATTGTTTGT	GCAGGGACT	CATAGACCA	29950
GGAAATGTCT	ATAGAGGTGG	GAATTTGTCT	CACCTGAAA	GGGATACCTC	3000

AAAGGGCTAT	TATCATTTTT	GTTCAGAGT	CAAACCATGA	ACTGAATTC	30800
TTCCCAAAGT	TAGTTCAGCC	TACACCCAGG	AATGAAGAAG	GACAGCTTAA	30850
AGGTTAGAAG	CAAGATGGAG	TCAATGAGGT	CTGATCTCTT	TCACTGTCAT	30900
AATTTCCCTCA	GTTATAATTT	TTGCAAAGGC	GGTTTCAGTC	CCAGCTACTT	30950
GGGAGGCTGA	GACAGGAGGA	TTAATGGAGC	CCAGGAGTTT	GAGGTTGCAG	31000
AGAGCTATGA	TCACGCCACT	GCACTCCAGC	CTGGGTGACA	GAGTGAGACC	31050
CTGTCTCTAA	ATAAATAAAT	AAGTAAATAA	ATAAATACAT	AAATAAAATC	31100
AAGATGGTGT	GCAATTAGAA	TTGAGCGATT	TTGTTTCCAA	ACCTCAAGAA	31150
AGCTTGGTCT	TGCTCTGTCC	CAGGTGGCTG	GATAAATTGG	GCCTGTGAGC	31200
CCGAATGGGA	ATAGAAGTGG	TGATGAGGCA	AGTATTCTTT	GGAGCAGGAA	31250
ACTACCATTT	AGTGGATGAA	AACTTCGATC	CTTTACCTGT	AAGTGACCAT	31300
TATTTTCCTA	ATTCTAGTGG	AGTAGATTAA	AGTCAACTCA	GGACCTCTGG	31350
TGTTAACCTC	CTATGAACAG	TCAGTCCCTC	CAGTAACTAG	CCAAATCATG	31400
AGATGATGAA	TTAGAAGGAG	CCTTAGATAG	CATCCAATCT	AACATTTTTT	31450
TGTGTGTTTG	AAGAGAAGAA	ATCAAGAGCT	AGGAATAACT	TTTTAAAGGT	31500
AAGCCATTTG	CAGTATAGTG	TGGATTTTGT	TTAAAGGGGG	ATAATTTGAA	31550
ATTTTATGAC	TCATTATACA	AGACAAAATA	AGTTGGATT	TCAAATGTTT	31600
TACAAAGTAA	ATCAAAGTTA	TAATTGCCTA	CAGTACGCAA	AGCTTCAAAA	31650
CATTTTTTAT	GTTATGAAAT	TGTAATTTAT	TTAACCTTAA	AATGAGCCAG	31700
TACCATGTGT	TTGCTTAAAA	ATCTCATGCT	AAGAATTTAC	TATGTTGTTA	31750
ATAATCTTCA	AGATATTTAT	GAATAAAGTC	TTATTCTTAA	TCCTTCCTCC	31800
AACTGTATCT	GGTGCTAAAT	CAGGAAATGT	TTCTTCCCAA	AAAGCCTCGT	31850
GGAAGATCTG	TATGTCTAAA	TATATGTCAG	GGATAATACA	GATGTAGCCC	31900
TGCGAAGCAT	GACCTTGATT	TTTATAGTCT	AAAATGTCAT	TTGCAGATAT	31950
CTATTTTCTA	AGAATAATTC	CTAAAAGAAT	TATTTGAATG	TTGTAGGAAA	32000
GCTAAGAAAT	TTTGCAAAGA	GCGTACGTGA	AAATATAAGC	TAGGCTTTTG	32050
TGGTTTGTGG	ATAGACTTCC	CAACAAAATT	GCTTTTTATC	TATAGTGATC	32100
CAAGCTTGTG	GAACATATTA	GTCATCTTTT	TTTAGAAAAA	TCTTAGAAAA	32150
GTGATCTTGC	AAAATGGAA	TTTATCTTTC	CCCAAGTATA	TTCTGTCATG	32200
TATAGAGTTA	AACTAAGCAT	AGTAATTTCA	CCAGACAAAC	ATTCAAAATC	32250
TACTCCTGAC	CTPTTTATCT	CATCCAAATT	TTCCAGGGG	CCAGACATAA	32300
ACCTTTGCCT	TACGAACCTC	TTGTATATGC	ACTAAATATG	CTTCTCCTTC	32350
AAGGTTCTCA	GTCAGCTAGA	AAAATGTGCA	AGAGTAAATG	GTACCTTTCT	32400
CACCTGTAGA	TCCAAGAGAA	TTAGACTTAA	ACTCACTCTA	CATGTCTGTG	32450
ACTTTATTTT	ATTTCATGTA	CAGTCCCTGT	AGGTGGCAAG	GCAGGTATCT	32500
TGGATCCATT	TTTTAGATAA	GGAAGTTCAA	ATTGAGAAGA	GGTTGCATGA	32550
TTTACAGGAA	GCCATACCTG	AGTCTATGTT	TACTCTTAAA	AATCCCATTC	32600
AAATCCTGCT	TCTGAGGCC	GCATACCTTC	TACCCATCCA	GTCATTGACC	32650
CATGCTTATG	TCTCCTTTGA	AAACATTGAT	TCCACTCTTG	TCTCCAGTGA	32700
AAAAGTGGAA	TTTTAAGCAGA	GAAACAAAAG	CCATTGTGCT	TGTTAAGTCT	32750
ACTTTCCCTC	TACTTTCAAG	AAGGAAAGTT	GGGGTATGTG	TTGAATGGTG	32800
ATTTATTTAT	TTATTTATTA	TTTTAAAAAT	TGATACAAGG	TCTTACTGTA	32850
TTGTGACGGC	TGGTCTCAAA	CTCCTGGGCT	CAAGTGATCA	TCCACCTCA	32900
GCCTCCAGT	GTTGGGATTA	CAGCATGAAC	CATTGTGCCC	ACCACCGATC	32950
CGCAGTTTTT	TAAGAAAAAC	TTTTACTATA	GAAAATTTTA	ATCATATACA	33000
AAATACAGAG	GAAAGTATAT	GAACCCACTT	TAGGAGACTA	GAATATGCCA	33050
CCCCAAATA	TGCCACTTTG	GCATAAGGAT	TATTTGAGC	TAAAGGCAAC	33100
TGGGAAGAAA	CACATAGAAG	AAAAGTTCTC	TGTCCTTCTC	CATTGCGCTA	33150
AAAGCAGGAC	ATGAATCTTA	AAAGTCCCC	TCCTTCCCTT	TCTACCAGGA	33200
AAAACAGAG	TTAATCACTG	AAGATAACTT	CAGACCCTTA	TCAGTGATGA	33250
GATGGCACTA	GAAGAATCTA	TATTACATAC	TCATTTATTT	TCCTTCCAC	33300
AACTTGCCAC	CCCAGAGACT	AAAAATCCTT	TTCTTTGTG	ATGTCTCTTG	33350
TCCAAAATTT	TGCTCTATAA	GCTGGAGTTC	TAAAGCCACT	CTTTGAGAA	33400
TACTTGTCTC	CTGGTATTTT	CTGTTAACAT	ACATGTATTA	ATATACATGT	33450
TAACAAGCTT	CTGTTTGTG	TTCTCCTGTT	TTCTGTCTTG	TTACAGAGGT	33500
CCATCCCAAC	TAAGAACTAA	AGAGTAGGAG	GAAAATATAA	TTTCTCTCTG	33550
CATACCTTGA	TCTTGTTTAA	TCCGTAACCC	TTCCCACCTT	TCACCTCCTA	33600
CCTATTAGAT	TACTTTGAAG	CAAATTTTCA	ATATATTACT	TTATCTATAA	33650
ATATTTTCA	ATGTGCTAGG	TGTGGTGGCT	CACACCTGTA	ATCCCAACAC	33700
TTTGGGAAGC	TGAGGCAGGA	GGATCACTTG	AGCCCAGGAG	TTCAAGACCA	33750
GCTACGGCAA	CAAAAAATCA	AAAACCTTAT	TGGGCATGGT	GGCACATGCC	33800
TGTGGTCCCA	GCTACATGAG	AGGCTGAGGC	AGGAGGATCG	CTTTAGCCCC	33850
GGAGGTTGAG	GCTGCAGTAA	GCTGCATTCA	CACCACTGCA	CTCCAGCCTG	33900
GGTGACAGAG	TAAGACCATG	TCTCAAAAAA	ATACATATTT	TAGTATGTAT	33950
CCTTTTGTGA	AAAACACAAT	ACTTTTATCA	TACTTTAAAT	AATAACAATA	34000
ATTCCTTAGT	ATCCCAAAAT	ATTTTGTGCT	TGTCCTCAT	TTTCTTATT	34050
GTCATAAATA	TTGTTGATAG	TTATTCAAAT	CAGAATCCAA	ACAAGGTCCA	34100
TATATTACAT	TTGGTTGACA	AGTCTCTTAA	GTTTGTTCAT	CTTTAAGTTC	34150
TTCTTCCCTC	TCTTTTCATCT	CTTGTAATTT	ATTAATGTGA	AAAAACAGGT	34200
AAATTTGTTT	ATAGTATTTT	CTACATTATA	GAGTTTGTCT	CATTTATTC	34250
CTATGATATC	ATTTAGCATG	TTCTCTGTG	CCCTGTGTTT	CCTGTAAACT	34300
GGTAGTTATA	CCTAGAAGCT	TGAGTTTATT	CAGGTTTTTA	ATTGTATTTT	34350
TTTTGCAAGA	ATTCTTTATT	ATCTGCTTCT	GGAAGCACAG	AATGTCTGGT	34400
TGTGTCTGGT	TTTGATCTTG	ACAGCTACTG	ATGACCATTG	CCTAATCCAT	34450
TACTTTATTG	GGGTGGGGGG	AATAAGGTTT	TAAATAAAAT	TTTTTTTAAA	34500
GATTTTTTTA	ACTGTTATTT	TGAGACAGTG	TCTCATTTCG	TTTCCAGGC	34550
TGGAGTGAG	TGGCACAATC	ACGGCTCACT	GCAGCCTTGA	CCTCCTGGGA	34600
TCAGGTGATC	TTCTCACCTC	AGCCCTCTGG	GTACCTGGAA	CTACAGGTGC	34650
ACACCACAC	ACCTGGCTAA	TTTTTGTGAT	TTTGTGTACA	GAAGGGGTTT	34700
CATCATGTTT	CCGAGACTGG	TCTTGAACCT	CTGGGTTCAA	GTGATCTACC	34750

CACCTTCAGCT	TCCCAAAATC	CTGGGATTAC	ACTTTGGCCA	CCGTGCCTGG	34800
CCTAAATGAA	ATTATTTGTC	TCTAAACAGA	CAGAAGTTTT	ACTTTAAAAA	34850
TTTGTCTTTG	TGTTGACATG	TGTTTGTGTA	TGTTGTGTGT	TCTAAAGATT	34900
TGGCTTTTGG	CTTTTCTTGG	AATTTCTTGA	TGAACAAATC	CCAGAAGATT	34950
TTAAACTCTG	ATCATTCTTG	ACAGATATCC	CCTACAGGCT	ATGGCCTTTT	35000
GAATTGTGTC	CTCCAGTGAT	AAAAAGCAGC	AAGACAGATA	CTGCTCTCAG	35050
ATTTCATGGT	GTGCACATGG	AGGTGAAAAA	AAAAAAAAG	ATGAACTCCTA	35100
TTTAAATGCC	CCCAGGATAA	CAGTGATAC	CTTTGTAGGA	TAACATTATTG	35150
CTTGCCCATG	GTTTTCATTAA	ATAAGGACAT	AAGTAAAGAT	CTATTTTTTG	35200
CTCTTTCTCC	CCAACCACCA	CAACTAGGAT	TATTGGCTAT	CTCTTCTGTT	35250
CAAGAAATTG	GTGGGCACCA	AGGTGTTAAT	GGCAAGCGTG	CAAGGTTCAA	35300
AGAGAAAGGA	CTCTCGAGTA	TACCTTCATT	GCACAAACAC	TGACAGTAA	35350
GTATGAAGAA	GACCTTTTAC	CATCATCAAA	GTTTTAGTGG	GTAAGCCTGT	35400
AACCTTTACTC	AAACACCCCTG	TTGCAATGTG	CTATACATTG	CATAGAGTATA	35450
GGCAGTTGCA	ATTTAGTAAA	GTTTTATACA	ACGATTTTAT	TTTATTTTAT	35500
TTTTAGAAGA	AAAATGCTAC	TTTTTGTGTT	GTGTTTTTTT	GAGACGGGGC	35550
CTCGCTCGTC	ACCCAGGGCTG	GAGTGCAGTG	GTGCAATCTC	AGCTCACTGC	35600
AACTCCCGCC	TCCCGGGTTC	AAGTGAATCT	GCACAAACAC	TGACAGTAA	35650
ACAACAATAT	TATTTTCAAA	AGTTGTGACC	TGAGTTTCTG	GAGTTGAGAA	35700
GACATCGAGA	TTTTTGTAGC	CTCATACTCT	TGCTTTAGGT	AGCAAAAAAT	35750
GTTCCTAAAT	CTACAGGAAT	TTCTCTAGAT	AGGTTTCAAT	CTATCATCTC	35800
TGTAAGATAG	ATGCTGAAT	ACTAATTTCTA	GCCAAAAAAG	ACCAGCTACC	35850
ATTTCCGATT	GTGCGGAGT	GGGAACCTCG	GATAGTGAGG	ACCCAGTAGC	35900
GAGTAGCGGA	GGGGAATGGT	TGTAATGGAT	AAATTCATAA	AAAATGTCAG	35950
TAGATTAAAT	TTTCTTATAC	ATTTCACTCT	TTTTATAAGG	CTAGGAAAAG	36000
CCCTCTGTTT	TATGGTTTAT	AAATTGAAAT	CACATGAACC	CAAAAAATTT	36050
GCCTTTTACC	TTCTTATGTC	TGAAGTAATG	TAGTCTGGCT	GGCCTCTTAA	36100
CAACCCGAGT	GGCAGAGCTG	TGAGGATCTC	AGTGTGCTCT	AGCCCAGACA	36150
TTGGTAGCAT	GAAACGCAAC	ATTTTAAAT	GTGTTTTCAA	AATGAGGACA	36200
CACATAGCGGT	CTAAAACGAT	CATAAAGAA	GGATACTAAG	AGGGCCACTC	36250
GTCAATTATG	ATTCCTAATAC	TTAGGATGCA	TTATGGATTG	TCATTATGGA	36300
TACTAATACT	TAGGATCACA	TTTGTAATGT	AGTTTTTAAT	TGCTTAAAT	36350
AGATACATAT	TTCTATTAA	TTTAACCTTT	TGCTTTTATG	CCAAGGTATA	36400
AAGAAGGAGA	TTTAACTCTG	TATGCCATAA	ACCTCCATAA	TGTCACCAAG	36450
TACTTTCGGT	TACCCCTATC	TTTTTCTAAC	AAGCAAGTGG	ATAAATACCT	36500
TCTAAGACCT	TTGGGACCTC	ATGGATATCT	TTCCAAGTAA	GTAATTTTTC	36550
TTGTTTCATC	CAACACTTCA	ATAAATTTAT	TGGTGTTTAT	CAGAATAGAG	36600
AGTTTGGACA	GGGAGCAAAA	GACAAAGTCA	ACTATATCAA	GTTCCTAATA	36650
TTCTTAATAT	TCAAGAAATT	TATGATGAA	TACTACTTAA	TATGAGTATA	36700
ACTCATCCTA	AGAGTCTAAA	GCAAAGGAT	GTGAACACAA	ACTAGCAGTT	36750
ATCTTAGAGA	ATAAGTTTGC	ATTTCAAAAT	AACCTTGACAT	ATCAAGATCC	36800
ACTCAACGCA	TTTAAATFAT	TTACTCTAAA	AAGACATAAT	TCTTGGTAAC	36850
ACATTCACTA	AAGCAAAATA	TACCTTTTAT	TAATTGCTAT	CAAAGGATAG	36900
TGGGTTGGTA	TAAATATATCA	TACCATTGTA	GTGCAGTGTG	ATTCCTTTAC	36950
AGCATTAAAT	TTTATTGGTT	AGAGTAAGAA	AAAGAATAGC	TAGAGTATAT	37000
TTCTTAAAGTA	GATTCTCATA	CACTTTGGTT	TCAAAAACCA	ATATTAGACT	37050
ACATCTTATA	AAAGCCTGTA	TTCAATGGAG	TGCCAAAAAA	TGCATTTGAG	37100
TCCTTAAAGAG	TTAGGCATAT	AAATATTTTA	AGGTTTCTGT	TCAATGTATG	37150
TTGGAAGGAG	TTCTTTTCTC	ATGACTATTC	TCATATTGGA	GCATAAAAAA	37200
AGTTTACAGG	CTTGGCGCAG	TGGCTCATGC	CTGTAATCCC	AATACTTTGG	37250
GAACTGTGAAG	AGGCAGATC	ACTTCAGCCC	AGGAGTTTGA	GACCCAGCCTG	37300
GGCAATATGG	CAAAACTCTC	TCTACAAAAA	ATACCAAAAT	TAGCCAGGCG	37350
TGGTGGTGCA	TGCCGTGTAG	CCCAAGCTACT	TGGAAGAGCTG	AGGTGGGAGG	37400
ATTGGCTTGAG	CCAGGGGGGG	TCAATGGCTG	AGTGAGCTGT	GATGGTGCCCT	37450
CTGTACACCA	GCCTGGGTGA	CAGAGTGAGA	CCCTGTCTCA	AAAAAATAAA	37500
TAAATAAAAA	TTAAGAGTTT	ACAAAATTTCT	CCCATCTCCC	TCCCATCTTT	37550
GCAAATGCCA	CATAAGTGAT	GTGTTCCAGT	ACTATTAGCC	TCGGAACCTG	37600
AGGCAGTACA	GTAAGCACGC	TTTCTCCAAA	GTCTGTCCCC	CCACAGACAA	37650
ACATTATTATA	CACCTGGGAT	TGCTCTTTTA	TTTTTCCCC	CTATGCTTTT	37700
ATTTTACTAT	AACTATAATC	ATATAACATG	TAATAGGAAA	AAGGCAGGGT	37750
CGGGGAGAG	ATCCAGAAGT	CTTCCCAAGA	GCCTTTCCAA	CATAGCCTCT	37800
GTAGACATTT	TTTCTTTCTT	CTTTTTTTTT	TTTTTTTTTT	TTCTGAGACA	37850
GAGTCTCACT	CTGTTGTCCA	GGCTAGAGTG	CAGTGGCGTG	ATCTAGGCTC	37900
ACTGCAACCT	CCGCTCCCTG	GGTTCAAGCA	ATTTCTCCAC	CTCAGCCTCC	37950
CTAGTAGCTG	GGATTAGAGT	CATGCATCAC	CACGCTGGC	TAATTTTTGT	38000
ATTTTTAGTA	GAGATGAGGT	TTACCAATGT	GGGCCAGGCT	GGTCTGAAC	38050
TCTGTACCTC	AAGCATATCA	CTGCTCTTAG	CTTCCCAAAG	TGCTAGGATT	38100
ACACGAGTGA	GGCACCCTGC	CCTGCCCTTA	TCATATTCTG	ATGCACATT	38150
TCATGTTTTA	TAATTTGAAA	ACTGGTGAAA	TTATAGACAA	TGTTTTGTTT	38200
CCCTAAATTT	TCCTTTGATG	GTATATATTA	CTTCACTCTT	TCTGTCTTTA	38250
AAATTTTGCA	AAATAGTATG	CTAGATAAGT	TATGAGTGTG	ACAGTCTGTA	38300
CGCTTACTCA	TATTAATGAC	CTCCGAGAGT	TAAACAACAG	TCACCTTTAA	38350
AAATATTATC	TATCATTTAT	ATTTATTTTT	AGGCGGGGGT	CTCATCTCTG	38400
CTCCAGGCT	GGAGATAGT	GGTGGGGTCA	CAGCTCACTG	CAGCCACCGC	38450
TACCTGGGCT	CAAGTGATCC	TTCTCTCTCA	GCCTTCTGAG	TAGCTGAGAC	38500
CACAGGCTTA	TGCTACACCA	CCTGGCTAAT	TTTTTAACTT	TTTGTAGAGA	38550
CGATGCTCTA	TATGTTGCC	CAGGCTGGTC	TCAAACCTCT	AAGTCAAGT	38600
GATCTTCCCT	AGCCTCCCAA	AGCTGCTGGG	TTACAGAGAT	GGAAAACTGC	38650
ACCCAGCCCT	AAAAATTATT	AGGTCCTGCA	ATAGTAGAGC	TTTTAATAAT	38700
ATTTAAATGA	CAACTTGGTT	TTTTTAAAAA	AAAAATAGAG	ACAAGGTCTC	38750

ACATATATTGC	CCAAGCTGGT	CTCGAACTCC	TGGACTCAGC	CAATCCTGCT	38800
GCCTTAGCCG	CCCAAAGTGC	TGGGATTACA	GGCATGACCC	ACCTCATCTG	38850
GGCTGAGTGA	ACATATTTTT	AACATAAAGG	CCGTATTTTA	TATTTATPCT	38900
ATACATTTTT	CCCAGCATAT	CCATTTCGCG	CGAATTCTGT	CTCTGCTAAT	38950
TCCTTCATCG	TTCATTTTCAT	CTGAAATTTG	ACAAACATCT	TCATTTTCTT	39000
TGTCGCTCAG	TTATTGACTT	CAGAATAATA	AGTAAAAAC	TATACCCAAA	39050
TTAAACCCEA	CCCTCATTTG	CCAGCCTGAT	ATGAAAAATA	TCAGCATACA	39100
TTAAGCTTAC	CCTTGATATA	TGTTGAGCAT	CTTTTAGATA	AATATACAGC	39150
CTTTTAAGCA	ATATAGCCTG	ATGGTATAAT	ATCTTGCCCA	TGACCTCAT	39200
CTTATPCTCA	GCAGGATTAA	TTCCACAGTA	TCAGATTTAC	CTTTAAACTT	39250
TGTAGCAAAA	TATCCTCTCC	AAAAGCATAT	CTAAAACTAT	TGTFGTGATC	39300
CTTGCAAGTT	TCTTAATTTT	ATGCGAACAA	GGCTCTTACC	ACTGTTAGCT	39350
GGAGATATTT	TCAGAACCTA	TTTTTGTTTG	TGGTTTCTGT	ATGATGGTCA	39400
TGGCATT TTC	CCCTTCACTC	CATCTAAAAA	TTGAGGTGAT	ACAGGCTTTT	39450
AAACAAAAAC	AACTCATATA	GACTGAGTAC	AAGTGCATAT	CAGGCATGCT	39500
AACCTCTGCT	ACAATCATGG	CGCTGCTATT	GATATGTCTT	AAGTTACAGA	39550
ACACAGGGCT	GAGCGTCTCA	TTAGGTCAAA	ATGATAAACCA	GTPTTTCTGC	39600
TCFACTGATG	TTAATGAGGA	CAGGGTGTGA	GAGATTTCCT	TAAGGAAAAAC	39650
AAATATAATA	TAATGCTACA	TGGAAAAATA	TCTAACATTA	GAGAATTAAG	39700
TAAATAAACT	AAATATACFA	CACCATGGAA	CTTTGTGCAG	AGATATAAAT	39750
TATGTAGTGG	ATGGATGTTT	ATATGTTGTGA	GAAAAAGTTA	GAGTGTGCTG	39800
GGGTGGGGGG	AAGAATCAAG	TTTTAAGAAA	ATACAGTATA	CCCATACTTA	39850
AGTAAAAAAA	AAAAAAAAGG	TATGTACAGT	CATGTGTTCG	TTATGTAGG	39900
GGATACATTC	CGAGAAATGT	GTCGATAGGT	GATTTTCATC	TTTGTGTCAAC	39950
ATCATAGAGT	GAACCTACAC	AAACCTAGAT	GGTCTAGCCT	ACTATGTATC	40000
TAGGCTATAT	GACTAGCCTG	TGCTCTCTAG	GCTACAAACC	TTAAAGCAT	40050
GTTACTGTAG	CGAATATACA	AATACTTAAC	ACAAATGGCAA	GCTATCATTT	40100
TGTTAAGTAG	TTGTGTATCT	AAACATATCT	AAAACATAGA	AAACTAATGT	40150
GTTGTGCTAC	AAATGTTACAA	TGACTATGAC	ATTGTCTAGC	AATAGGAATT	40200
ATAAATTTAT	CCTTTATATG	ACCACCATCT	ATATATGCGG	TCCATGTGGG	40250
ACCAAAAATC	CCTTATGTGG	CATATGACTG	TATACATGTA	CACAAAAAAT	40300
AGATGAAGAA	ATGAATATAC	ATCAAAATAT	TTAAAGTTGT	TATAATGACT	40350
TAGGTTTACT	TTTATATCT	TAGTAATAAT	ATAGATGATA	GATAATACCT	40400
TTATAGTGTT	TACTATATAA	AAGACACTGT	TATAAGTGTT	CTACATACTT	40450
TACATGTATT	ACCTAAATGA	TATAAATATA	ACTCTGACAG	TAACATTAAT	40500
TATACGTCTT	CTTTTCTTTT	TTTTTTTTTT	CTTTTTTTAG	ACAGAACTCT	40550
GCTCTACCAG	GCTGGAGTGC	AGGGTGCAAT	CTCGGCTCAC	TGCAACCTCC	40600
GCCTCCCGAG	TTCAAACGAT	TCTCATGTCT	CAGCCTCCGT	AGTAGCTGGG	40650
ACTACGAGCA	CACACACCCA	GGCCCGGCTA	ATTTTTGTAT	TTTTGGGTAG	40700
AGATGGAGTT	TTGCCATGTT	GGCCAGGCTG	ATCTTGAATT	CCTGGCCTCA	40750
ATPGATCTGC	GCTGCTCAGC	CTCCCAAAGT	GCTGGGATTA	CAGGTGTGAA	40800
CCACTGTGCT	CGGCTTAATC	TTACAAGTTT	TCBAATATTA	AAGAGTGCTA	40850
ACTTTGTTGA	CAATATAAAA	CATATTTGAG	AAAAAGAGAT	ATAGCATCTT	40900
TTTTTAGAAT	TATGAAAATA	TCAATAGACC	TACAGCCGAC	TAAAGCTTTT	40950
CTTATATAAG	TCTTGCCATAT	ATTGATTCGC	CTCTGTGAAT	ATGCATTAAT	41000
TTGATTTAAA	TAATAAGTAT	GTATAAGAAA	TAACACTTTT	CCTTAATTTT	41050
TAAGAACGTT	CAACAGCTTT	TAATTTGAAT	TCCAATAGTG	AAATACATAG	41100
AAAAATATAA	ATTTTCTGTA	GTTTAGCCAA	ATTGTTTTTG	TTTACCACAA	41150
GCATTCTACC	AAAATTTCTT	AATAACAGTA	AGAAAAATGA	TGCATACCTC	41200
CTCAGGGGAG	AGGGGAGTTA	GGCAGTTTAT	GGGCATAGTT	ACAAGTGAGA	41250
AAATTTCACT	GCTACCAATT	ACGCTAAATC	CATAAAAACT	GACTTCAATT	41300
CTATATATCT	ATTTTCTTTA	CATAAAAAAG	TTTCAATTA	TTGGCCATTA	41350
ATAAAAAATG	CCCAACTTCC	AGAAGTTTGT	TCTGTTTTAT	CCTTTTTATA	41400
CCACCATCAT	ATTGCTTATT	ATATAGATTG	TGTTGTTTCC	ATTTTCTGTA	41450
ATGGGCCAGA	CAGTAAGTAT	TTCTGGCTTT	GGAGTCCATA	TGGTCTCTAT	41500
CATAACTACT	CATCTCTGCC	ATTGTAGCTT	AAAGATTATC	TAGGTCAAAT	41550
GCCTAAGTGA	TATAGTGTTG	AAATACAAGT	TATATAATAT	AGGCTGCCAC	41600
AAAAAAAATAT	TTATTTGGTC	TAAAAAAGAT	TTTCATGACT	TTGTAGCAGC	41650
ATGGGTGGGG	CATGCCAACG	TTGGTTAACT	CCGTGATCTT	TCTCTCTTTG	41700
GAGGCTGTCT	CAACTCAATG	GTCTAACTCT	AAAGATGGTG	GATGATCAAA	41750
CCTTGCCACC	TTTAATGGAA	AAACCTCTCC	GGCCAGGAAG	TTCACTGGGC	41800
TGTCGACGTT	TCTCATATAG	TTTTTTTGTG	ATAAGAAATG	CCAAAGTTGC	41850
TGCTTGCAAT	TGAAAAATAA	ATATACTAGT	CCTGACACTG	AATTTTTCAA	41900
GTATFACTAG	AGTAAAGCAA	CTCAAGTTAT	AGGAAAGGAA	GCAGATACCT	41950
TGCAAGCCAA	CAGTAGGGTG	CTTGAGAGAC	ACTGGGACAC	TGTCAGTGCT	42

CTCTGTGCAG TGTGTCTAGT CAGCGAAAGA CTATAATACC TGTGGGGACA 42800
 GCGATTAGCC ACCACAACCA GTCTTTATTT AAAGTTATTA AAAATGGCTG 42850
 GGCGCAGTGG CTCACACCTG TAATCCTAGC ACTTTGGGAG GCCGAGGCAG 42900
 ATGGATCACC TGACGTGAGG AATTTGAGAC CAGCCTGGCC AACATGGTGA 42950
 AACCCCATCT CTACTAAAAA ATACAAAAAT TAGCTGGGTG TGGTCTCTGA 43000
 GTCCAGCTA CTGTTGGGAG TGGGGCAGGA GAATTACTTG AACCCAGGAG 43050
 GCAGAGGTTG CAGTGAGCCG AGATTGTGCC ACTGCACTCC AGCCTGGGTG 43100
 ACAGAGAGAG ATTCCATCTC AAAAAAACAA GTTATTAAAA ATGTATATGA 43150
 ATGCTCCTAA TATGGTCAGG AAGCAAGGAA GCGAAGGATA TATTATGAGT 43200
 TTTAAGAAGG TGCTTAGCTG TATATTATC TTTCAAAATG TATTAGAAGA 43250
 TTTTAGAATT CTTTCCTTCA TGTGCCATCT CTACAGGCAC CCATCAGAAA 43300
 AAGCATACTG CCGTTACCGT GAAACTGGTT GTAAAAAGAGA AACTATCTAT 43350
 TTGCACCTTA AAAGACAGCT AGATTTTGCT GATTTTCTTC TTTCGGTTTT 43400
 CTTTGTGCAG AATAATATGT GAGAGGACAG ATTGTTAGAT ATGATAGTAT 43450
 AAAAAATGGT TAATGACAAAT TCAGAGGCGA GGAGATTCTG TAAACTTAAA 43500
 ATTACTATAA ATGAAATTGA TTTGTCAAGA GGATAAATTT TAGAAAAACAC 43550
 CCAATACCTT ATAAGTGTCT GTTAATGCTT GCTTTTCTC TACCTTCTT 43600
 CCTTGTTCCT GTTGGGAGG TTTTGGCTGC AAGTAACAGA AACTCCTAAT 43650
 TCAAAATGGT TAAGCAATAA GGAAATGTAT ATTCCACAT AACTAGACGT 43700
 TCAAAACAGG CAGGCTCCAG CACTTCAGTA CGTCACCAGG GATCTGGGT 43750
 CTTCCAGCT CTCTGCTCTG CCATCTTTAG CGCTGGCTTC ATTCTCAGAC 43800
 TCTGTAGCA TGATGGCTGT AGCTGTTTCA TGGGCCCTT CAAACCTCAT 43850
 AGCAACCAGA GGAAGAAAT GAGCCATTTT TTGAGTCTCC TTCATAGACT 43900
 TGAATAACTC TTTTTCAGAG CTCTCTCAG CAAACCTCTC CTCATGTCTC 43950
 CTCATGTCTT ATGTTTCAAGA AATGGGTAAT GTGGCCATTT CACCAGTCAC 44000
 TGCCAAACAC AACGAGGTTT CTATAATTGT CTCTGAGTAA CCCTTGGAA 44050
 TGGAGAGGGT GTTGGTCAGT CTACAACTG AACACTGCAG TTCTGCGCTT 44100
 TTTACCAGTG AAAAAATGTA ATTATTTTCC CCTCTTAAGG ATTAATATT 44150
 TTCAAATGTA TGCTGTATAT GGATATAGTA TCTTTAAAT TTTTATTTT 44200
 AATAGCTTTA GGGGTACACA CTTTTTGCTT ACAGGGGTGA ATTGTGTAGT 44250
 GGTGAAGACT CGGCTTTTAA TGTACTTGT ACCTGAGTGA TGTACATTGT 44300
 ACCCAATAGG TAATTTTCA TCCATTACCC TCCTTCCGCC CTCTTCCCTT 44350
 CTGAGTCTCC AACATCCCTT ATACCCTGT GTATGTTCTT GTGTACCTAC 44400
 AGCTAAGCTT CCACCTTATA GTGAGAACAT GCAGTATTTG GTTTTCCATT 44450
 CCTGAGTTAC TTCCCTTAGG ATAACAGCCC CAGTTCCGT CCAAGTTGCT 44500
 GCAAAATACA TTATTCTTCT TTATGGCTGA GTAATAGTCC ATGGTACATA 44550
 TATACCACAT TTTCTTTATC CACTTATCAG TTGATGGACA CTTAGGTTAA 44600
 TTCCATTCAA TTTTATTCAA TTTAAGTATA TTTGTAAGGA GCTAAAGCTG 44650
 AAAATTAAAT TTTAGATCTT TCAATACTCT TAAATTTTAT ATGTAAGTGG 44700
 TTTTATATTT TTACATTTG AATAAAGTA ATTTTATATA CCTTGATATT 44750
 GTATGACTAT TCTTTAGTA ATGTAAAGCC TACAGACTCC TACATTGGA 44800
 ACCACTAGTG TGTGTTTCA CCCCTTGTTA TACTATCAGG ATCCTCGA 44848

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2396
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43

TTTCTAGTTG CTTTtagcca ATGTCGGATC AGGTTTTTCA AGCGACAAAG 50
 AGATACTGAG ATCCTGGGCA GAGGACATCC TAGCTCGGTC AGATTGGGGC 100
 AGGCTCAAGT GACCAGTGTC TTAAGGCAGA AGGGAGTCGG GGTAGGGTCT 150
 GGCTGAACCC TCAACCGGGG CTTTAACTC AGGGTCTAGT CCTGGCGCCA 200
 AATGGATGGG ACCTAGAAAA GGTGACAGAG TGCAGCAGGAC ACCAGGAAGC 250
 TGGTCCCACC CCTGCGCGGC TCCCGGGCGC TCCCTCCCA GGCCTCCGAG 300
 GATCTTGGAT TCTGGCCACC TCCGCACCTT TTGGATGGGT GTGGATGATT 350
 TCAAAAGTGG ACGTGACCGC GCGGAGGGG AAAGCCAGCA CGGAAATGAA 400
 AGAGAGCGAG GAGGGGAGGG CGGGGAGGGG AGGGCGCTAG GGAGGGACTC 450
 CCGGGAGGGG TGGGAGGGAT GGAGCGCTGT GGGAGGGTAC TGAGTCCTGG 500
 CGCCAGAGGC GAAGCAGGAC CGGTTGCAGG GGGCTTGAGC CAGCGCGCCG 550
 GCTGCCCCAG CTCTCCCGGC AGCGGGCGGT CCAGCCAGGT GGGATGCTGA 600
 GGCTGCTGCT GCTGTGGCTC TGGGGGCGGC TCGGTGCCCT GGCCAGGGC 650
 GCCCCGCGG GGACCGCGCC GACCGACGAC GTGGTAGACT TGGAGTTTAA 700
 CACCAAGCGG CCGCTCCGAA GCGTGAGTCC CTCGTTCCCTG TCCATACCA 750
 TCGACGCCAG CCTGGCCACC GACCCGCGCT TCCTCACCTT CCTGGGCTCT 800
 CCAAGGCTCC GTGCTCTGGC TAGAGGCTTA TCTCTGCAT ACTTGAGATT 850
 TGGCGGCACA AAGACTGACT TCCTTATTTT TGATCCGGAC AAGGAACCGA 900
 CTTCCGAAGA AAGAAGTTAC TGGAAATCTC AAGTCAACCA TGATATTGCT 950

AGGTCTGAGC CGGTCTCTGC TGCGGTGTTG AGGAAACTCC AGGTGGAATG 1000
 GCCCTTCCAG GAGCTGTTGC TGCTCCGAGA GCAGTACCAA AAGGAGTTCA 1050
 AGAACAGCAC CTACTCAAGA AGCTCAGTGG ACATGCTCTA CAGTTTTGCC 1100
 AAGTGCTCGG GGTAGACCT GATCTTTGGT CTAAATGCGT TACTACGAAC 1150
 CCCAGACTTA CGGTGGAACA GcTCCAACGC CCAGCTTCTC CTTGACTACT 1200
 GCTCTTCCAA GGGTTATAAC ATcTCCTGGG AACTGGGCAA TGAGCCCAAC 1250
 AGTTTcTGGA AGAAAGCTCA CATTCTCATC GATGGGTTGC AGTTAGGAGA 1300
 AGACTTTGTG GAGTTGCATA AACTTcTACA AAGGTCAGCT TTCCAAAATG 1350
 CAAAACTCTA TGGTCTGAC ATCGGTCAGC CTCGAGGGAA GACAGTTAAA 1400
 CTGCTGAGGA GTTTCCTGAA GGCTGGCGGA GAAGTGATCG ACTCTCTTAC 1450
 ATGGCATCAC TATTACTTGA ATGGACGCAT CGCTACCAAA GAAGATTTTC 1500
 TGAGCTCTGA TGCGCTGGAC ACTTTTATTC TCTCTGTGCA AAAAATTCTG 1550
 AAGGTCACTA AAGAGATCAC ACCTGGCAAG AAGGTCTGGT TGGGAGAGAC 1600
 GAGCTCAGCT TACGGTGGCG GTGCACCCTT GCTGTCCAAC ACCTTTCAG 1650
 CTGGCTTTAT GTGGCTGGAT AAATTGGGCC TGTCAGCCCA GATGGGCATA 1700
 GAAGTCGTGA TGAGGCAGGT GTTCTTCGGA GCAGGCAACT ACCACTAGT 1750
 GGATGAAAAC TTTGAGCCTT TACCTGATTA CTGGCTCTCT CTTCTGTTCA 1800
 AGAAACTGGT AGGCTCCAGG GTGTTACTGT CAAGAGTGAA AGGCCAGAC 1850
 AGGAGCAAAG TCCGAGTGTA TCTCCACTGC ACTAACGTCT ATCACCACG 1900
 ATATCAGGAA GGAGATCTAA CTCTGTATGT CCTGAACCTC CATAATGTCA 1950
 CCAAGCACTT GAAGGTACCG CCTCCGTTGT TCAGGAAACC AGTGGATACG 2000
 TACCTTCTGA AGCCTTCGGG GCCGGATGGA TTACTTTCCA AATCTGTCCA 2050
 ACTGAACGGT CAAATTCTGA AGATGGTGGA TGAGCAGACC CTGCCAGCTT 2100
 TGACAGAAAA ACCTCTCCCC GCAGGAAGTG CACTAAGCCT GCCTGCCTTT 2150
 TCCTATGGTT TTTTGTGCAT AAGAAATGCC AAAATCGCTG CTTGTATATG 2200
 AAAATAAAAG GCATACGGTA CCCCTGAGAC AAAAGCCGAG GGGGGTGTTA 2250
 TTCATAAAAC AAAACCTAG TTTAGGAGGC CACCTCCTTG CCGAGTTCCA 2300
 GAGCTTCGGG AGGGTGGGGT ACACITCAGT ATTACATTCA GTGTGGTGTT 2350
 CTCTCTAAGA AGAATACTGC AGGTGGTGAC AGTTAATAGC ACTGTG 2396

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 535
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:44

Met Leu Arg Leu Leu Leu Trp Leu Trp Gly Pro Leu Gly Ala
 5 10 15
 Leu Ala Gln Gly Ala Pro Ala Gly Thr Ala Pro Thr Asp Asp Val
 20 25 30
 Val Asp Leu Glu Phe Tyr Thr Lys Arg Pro Leu Arg Ser Val Ser
 35 40 45
 Pro Ser Phe Leu Ser Ile Thr Ile Asp Ala Ser Leu Ala Thr Asp
 50 55 60
 Pro Arg Phe Leu Thr Phe Leu Gly Ser Pro Arg Leu Arg Ala Leu
 65 70 75
 Ala Arg Gly Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly Thr Lys
 80 85 90
 Thr Asp Phe Leu Ile Phe Asp Pro Asp Lys Glu Pro Thr Ser Glu
 95 100 105
 Glu Arg Ser Tyr Trp Lys Ser Gln Val Asn His Asp Ile Cys Arg
 110 115 120
 Ser Glu Pro Val Ser Ala Ala Val Leu Arg Lys Leu Gln Val Glu
 125 130 135
 Trp Pro Phe Gln Glu Leu Leu Leu Leu Arg Glu Gln Tyr Gln Lys
 140 145 150

Glu Phe Lys Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Met Leu	155	160	165
Tyr Ser Phe Ala Lys Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu	170	175	180
Asn Ala Leu Leu Arg Thr Pro Asp Leu Arg Trp Asn Ser Ser Asn	185	190	195
Ala Gln Leu Leu Leu Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile	200	205	210
Ser Trp Glu Leu Gly Asn Glu Pro Asn Ser Phe Trp Lys Lys Ala	215	220	225
His Ile Leu Ile Asp Gly Leu Gln Leu Gly Glu Asp Phe Val Glu	230	235	240
Leu His Lys Leu Leu Gln Arg Ser Ala Phe Gln Asn Ala Lys Leu	245	250	255
Tyr Gly Pro Asp Ile Gly Gln Pro Arg Gly Lys Thr Val Lys Leu	260	265	270
Leu Arg Ser Phe Leu Lys Ala Gly Gly Glu Val Ile Asp Ser Leu	275	280	285
Thr Trp His His Tyr Tyr Leu Asn Gly Arg Ile Ala Thr Lys Glu	290	295	300
Asp Phe Leu Ser Ser Asp Ala Leu Asp Thr Phe Ile Leu Ser Val	305	310	315
Gln Lys Ile Leu Lys Val Thr Lys Glu Ile Thr Pro Gly Lys Lys	320	325	330
Val Trp Leu Gly Glu Thr Ser Ser Ala Tyr Gly Gly Gly Ala Pro	335	340	345
Leu Leu Ser Asn Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys	350	355	360
Leu Gly Leu Ser Ala Gln Met Gly Ile Glu Val Val Met Arg Gln	365	370	375
Val Phe Phe Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe	380	385	390
Glu Pro Leu Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu	395	400	405
Val Gly Pro Arg Val Leu Leu Ser Arg Val Lys Gly Pro Asp Arg	410	415	420
Ser Lys Leu Arg Val Tyr Leu His Cys Thr Asn Val Tyr His Pro	425	430	435
Arg Tyr Gln Glu Gly Asp Leu Thr Leu Tyr Val Leu Asn Leu His	440	445	450
Asn Val Thr Lys His Leu Lys Val Pro Pro Pro Leu Phe Arg Lys	455	460	465
Pro Val Asp Thr Tyr Leu Leu Lys Pro Ser Gly Pro Asp Gly Leu	470	475	480
Leu Ser Lys Ser Val Gln Leu Asn Gly Gln Ile Leu Lys Met Val	485	490	495
Asp Glu Gln Thr Leu Pro Ala Leu Thr Glu Lys Pro Leu Pro Ala	500	505	510
Gly Ser Ala Leu Ser Leu Pro Ala Phe Ser Tyr Gly Phe Phe Val	515	520	525
Ile Arg Asn Ala Lys Ile Ala Ala Cys Ile	530	535	

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2396

(B) TYPE: nucleic acid

00881.13.14.01

(C) STRANDEDNESS: double
 (D) TOPOLOGY: linear
 (x1) SEQUENCE DESCRIPTION: SEQ ID NO:45

	TT TCT AGT	8
TGC TTT TAG CCA ATG TCG GAT CAG GTT TTT CAA GCG ACA AAG AGA		53
TAC TGA GAT CCT GGG CAG AGG ACA TCC TAG CTC GGT CAG ATT TGG		98
GCA GGC TCA AGT GAC CAG TGT CTT AAG GCA GAA GGG AGT CGG GGT		143
AGG GTC TGG CTG AAC CCT CAA CCG GGG CTT TTA ACT CAG GGT CTA		188
GTC CTG GCG CCA AAT GGA TGG GAC CTA GAA AAG GTG ACA GAG TGC		233
GCA GGA CAC CAG GAA GCT GGT CCC ACC CCT GCG CGG CTC CCG GGC		278
GCT CCC TCC CCA GGC CTC CGA GGA TCT TGG ATT CTG GCC ACC TCC		323
GCA CCC TTT GGA TGG GTG TGG ATG ATT TCA AAA GTG GAC GTG ACC		368
GCG GCG GAG GGG AAA GCC AGC ACG GAA ATG AAA GAG AGC GAG GAG		413
GGG AGG GCG GGG AGG GGA GGG CGC TAG GGA GGG ACT CCC GGG AGG		458
GGT GGG AGG GAT GGA GCG CTG TGG GAG GGT ACT GAG TCC TGG CGC		503
CAG AGG CGA AGC AGG ACC GGT TGC AGG GGG CTT GAG CCA GCG CGC		548
CGG CTG CCC CAG CTC TCC CGG CAG CGG GCG GTC CAG CCA GGT GGG		593
ATG CTG AGG CTG CTG CTG CTG TGG GGG CCG CTC GGT GCC		638
Met Leu Arg Leu Leu Leu Trp Leu Trp Gly Pro Leu Gly Ala		
5 10 15		
CTG GCC CAG GGC GCC CCC GCG GGG ACC GCG CCG ACC GAC GAC GTG	683	
Leu Ala Gln Gly Ala Pro Ala Gly Thr Ala Pro Thr Asp Asp Val		
20 25 30		
GTA GAC TTG GAG TTT TAC ACC AAG CGG CCG CTC CGA AGC GTG AGT	728	
Val Asp Leu Glu Phe Tyr Thr Lys Arg Pro Leu Arg Ser Val Ser		
35 40 45		
CCC TCG TTC CTG TCC ATC ACC ATC GAC GCC AGC CTG GCC ACC GAC	773	
Pro Ser Phe Leu Ser Ile Thr Ile Asp Ala Ser Leu Ala Thr Asp		
50 55 60		
CCG CGC TTC CTC ACC TTC CTG GGC TCT CCA AGG CTC CGT GCT CTG	818	
Pro Arg Phe Leu Thr Phe Leu Gly Ser Pro Arg Leu Arg Ala Leu		
65 70 75		
GCT AGA GGC TTA TCT CCT GCA TAC TTG AGA TTT GGC GGC ACA AAG	863	
Ala Arg Gly Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly Thr Lys		
80 85 90		
ACT GAC TTC CTT ATT TTT GAT CCG GAC AAG GAA CCG ACT TCC GAA	908	
Thr Asp Phe Leu Ile Phe Asp Pro Asp Lys Glu Pro Thr Ser Glu		
95 100 105		
GAA AGA AGT TAC TGG AAA TCT CAA GTC AAC CAT GAT ATT TGC AGG	953	
Glu Arg Ser Tyr Trp Lys Ser Gln Val Asn His Asp Ile Cys Arg		
110 115 120		
TCT GAG CCG GTC TCT GCT GCG GTG TTG AGG AAA CTC CAG GTG GAA	998	
Ser Glu Pro Val Ser Ala Ala Val Leu Arg Lys Leu Gln Val Glu		
125 130 135		
TGG CCC TTC CAG GAG CTG TTG CTG CTC CGA GAG CAG TAC CAA AAG	1043	
Trp Pro Phe Gln Glu Leu Leu Leu Arg Glu Gln Tyr Gln Lys		
140 145 150		

0998813 "11901

GAG	TTC	AAG	AAC	AGC	ACC	TAC	TCA	AGA	AGC	TCA	GTG	GAC	ATG	CTC	1088
Glu	Phe	Lys	Asn	Ser	Thr	Tyr	Ser	Arg	Ser	Ser	Val	Asp	Met	Leu	
				155					160					165	
TAC	AGT	TTT	GCC	AAG	TGC	TCG	GGG	TTA	GAC	CTG	ATC	TTT	GGT	CTA	1133
Tyr	Ser	Phe	Ala	Lys	Cys	Ser	Gly	Leu	Asp	Leu	Ile	Phe	Gly	Leu	
				170					175					180	
AAT	GCG	TTA	CTA	CGA	ACC	CCA	GAC	TTA	CGG	TGG	AAC	AGC	TCC	AAC	1178
Asn	Ala	Leu	Leu	Arg	Thr	Pro	Asp	Leu	Arg	Trp	Asn	Ser	Ser	Asn	
				185					190					195	
GCC	CAG	CTT	CTC	CTT	GAC	TAC	TGC	TCT	TCC	AAG	GGT	TAT	AAC	ATC	1223
Ala	Gln	Leu	Leu	Leu	Asp	Tyr	Cys	Ser	Ser	Lys	Gly	Tyr	Asn	Ile	
				200					205					210	
TCC	TGG	GAA	CTG	GGC	AAT	GAG	CCC	AAC	AGT	TTC	TGG	AAG	AAA	GCT	1268
Ser	Trp	Glu	Leu	Gly	Asn	Glu	Pro	Asn	Ser	Phe	Trp	Lys	Lys	Ala	
				215					220					225	
CAC	ATT	CTC	ATC	GAT	GGG	TTG	CAG	TTA	GGA	GAA	GAC	TTT	GTG	GAG	1313
His	Ile	Leu	Ile	Asp	Gly	Leu	Gln	Leu	Gly	Glu	Asp	Phe	Val	Glu	
				230					235					240	
TTG	CAT	AAA	CTT	CTA	CAA	AGG	TCA	GCT	TTC	CAA	AAT	GCA	AAA	CTC	1358
Leu	His	Lys	Leu	Leu	Gln	Arg	Ser	Ala	Phe	Gln	Asn	Ala	Lys	Leu	
				245					250					255	
TAT	GGT	CCT	GAC	ATC	GGT	CAG	CCT	CGA	GGG	AAG	ACA	GTT	AAA	CTG	1403
Tyr	Gly	Pro	Asp	Ile	Gly	Gln	Pro	Arg	Gly	Lys	Thr	Val	Lys	Leu	
				260					265					270	
CTG	AGG	AGT	TTC	CTG	AAG	GCT	GGC	GGA	GAA	GTG	ATC	GAC	TCT	CTT	1448
Leu	Arg	Ser	Phe	Leu	Lys	Ala	Gly	Gly	Glu	Val	Ile	Asp	Ser	Leu	
				275					280					285	
ACA	TGG	CAT	CAC	TAT	TAC	TTG	AAT	GGA	CGC	ATC	GCT	ACC	AAA	GAA	1493
Thr	Trp	His	His	Tyr	Tyr	Leu	Asn	Gly	Arg	Ile	Ala	Thr	Lys	Glu	
				290					295					300	
GAT	TTT	CTG	AGC	TCT	GAT	GCG	CTG	GAC	ACT	TTT	ATT	CTC	TCT	GTG	1538
Asp	Phe	Leu	Ser	Ser	Asp	Ala	Leu	Asp	Thr	Phe	Ile	Leu	Ser	Val	
				305					310					315	
CAA	AAA	ATT	CTG	AAG	GTC	ACT	AAA	GAG	ATC	ACA	CCT	GGC	AAG	AAG	1583
Gln	Lys	Ile	Leu	Lys	Val	Thr	Lys	Glu	Ile	Thr	Pro	Gly	Lys	Lys	
				320					325					330	
GTC	TGG	TTG	GGA	GAG	ACG	AGC	TCA	GCT	TAC	GGT	GGC	GGT	GCA	CCC	1628
Val	Trp	Leu	Gly	Glu	Thr	Ser	Ser	Ala	Tyr	Gly	Gly	Gly	Ala	Pro	
				335					340					345	
TTG	CTG	TCC	AAC	ACC	TTT	GCA	GCT	GGC	TTT	ATG	TGG	CTG	GAT	AAA	1673
Leu	Leu	Ser	Asn	Thr	Phe	Ala	Ala	Gly	Phe	Met	Trp	Leu	Asp	Lys	
				350					355					360	

TTG GGC CTG TCA GCC CAG ATG GGC ATA GAA GTC GTG ATG AGG CAG 1718
 Leu Gly Leu Ser Ala Gln Met Gly Ile Glu Val Val Met Arg Gln
 365 370 375

GTG TTC TTC GGA GCA GGC AAC TAC CAC TTA GTG GAT GAA AAC TTT 1763
 Val Phe Phe Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe
 380 385 390

GAG CCT TTA CCT GAT TAC TGG CTC TCT CTT CTG TTC AAG AAA CTG 1808
 Glu Pro Leu Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu
 395 400 405

GTA GGT CCC AGG GTG TTA CTG TCA AGA GTG AAA GGC CCA GAC AGG 1853
 Val Gly Pro Arg Val Leu Leu Ser Arg Val Lys Gly Pro Asp Arg
 410 415 420

AGC AAA CTC CGA GTG TAT CTC CAC TGC ACT AAC GTC TAT CAC CCA 1898
 Ser Lys Leu Arg Val Tyr Leu His Cys Thr Asn Val Tyr His Pro
 425 430 435

CGA TAT CAG GAA GGA GAT CTA ACT CTG TAT GTC CTG AAC CTC CAT 1943
 Arg Tyr Gln Glu Gly Asp Leu Thr Leu Tyr Val Leu Asn Leu His
 440 445 450

AAT GTC ACC AAG CAC TTG AAG GTA CCG CCT CCG TTG TTC AGG AAA 1988
 Asn Val Thr Lys His Leu Lys Val Pro Pro Pro Leu Phe Arg Lys
 455 460 465

CCA GTG GAT ACG TAC CTT CTG AAG CCT TCG GGG CCG GAT GGA TTA 2033
 Pro Val Asp Thr Tyr Leu Leu Lys Pro Ser Gly Pro Asp Gly Leu
 470 475 480

CTT TCC AAA TCT GTC CAA CTG AAC GGT CAA ATT CTG AAG ATG GTG 2078
 Leu Ser Lys Ser Val Gln Leu Asn Gly Gln Ile Leu Lys Met Val
 485 490 495

GAT GAG CAG ACC CTG CCA GCT TTG ACA GAA AAA CCT CTC CCC GCA 2123
 Asp Glu Gln Thr Leu Pro Ala Leu Thr Glu Lys Pro Leu Pro Ala
 500 505 510

GGA AGT GCA CTA AGC CTG CCT GCC TTT TCC TAT GGT TTT TTT GTC 2168
 Gly Ser Ala Leu Ser Leu Pro Ala Phe Ser Tyr Gly Phe Phe Val
 515 520 525

ATA AGA AAT GCC AAA ATC GCT GCT TGT ATA TGA AAA TAA AAG GCA 2213
 Ile Arg Asn Ala Lys Ile Ala Ala Cys Ile
 530 535

TAC GGT ACC CCT GAG ACA AAA GCC GAG GGG GGT GTT ATT CAT AAA 2258
 ACA AAA CCC TAG TTT AGG AGG CCA CCT CCT TGC CGA GTT CCA GAG 2303
 CTT CGG GAG GGT GGG GTA CAC TTC AGT ATT ACA TTC AGT GTG GTG 2348
 TTC TCT CTA AGA AGA ATA CTG CAG GTG GTG ACA GTT AAT AGC ACT 2393
 GTG 2396

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 385

(B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46

```
CGGCCGCTGC TGCTGCTGTG GCTCTGGGGG CGGCTCCGTG CCCTGACCCA 50
AGGCACTCCG GCGGGGACCG CGCCGACCAA AGACGTGGTG GACTTGGAGT 100
TTTACACCAA GAGGCTATTC CAAAGCGTGA GTCCCTCGTT CCTGTCCATC 150
ACCATCGACG CCAGTCTGGC CACCGACCCT CGGTTCTCTCA CCTTCCTGAG 200
CTCTCCACGG CTTCGAGCCC TGTCTAGAGG CTTATCTCCT GCGTACTTGA 250
GATTTGGCGG CACCAAGACT GACTTCCTTA TTTTGTATCC CAACAACGAA 300
CCCACCTCTG AAGAAAGAAG TTAAGTGGCA TCTCAAGACA ACAATGATAT 350
TTGCGGGTCT GACCGGGTCT CCGCTGACGT GTTGA 385
```

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 541
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47

```
AAATCAGGAC ATATCCTTCA CTTATTTGCC TCTTGGTCAT ATTGGAGGCA 50
TTTGTATTCA TTTTAAATAA CCCTCAAAAT AGTGCATGCA AAGTGCTAAG 100
CGTCATTTCG CACATGGTGC CATTAACTGT CACCACCTGC AGTGGTCTAC 150
TTAGAGAACA CCGCACTGGA TGTTAACTACT GAAGCGCGTG CCCC GCCCTC 200
CCGAGGCTCT GGATCCAGCG TTGAAGCTTG CCCC GCCCTC CCGAGGCTCT 250
GGATCCAGCA CTGGAGCATG CCCC GCCCTC CCGAGGCTCT GGAGCTTGCT 300
AAGGAGTCCG CTCCCTACCG CTGGGGTTTT GCTTTATCT TATGAATGAC 350
ACCCCTGACC GCTTTCGTCT CAGGGGTACT GTAATGCCTT TTATTTTCAT 400
ATACAAGCTG CGATTTTGGC ATTTCTTATG ACAAAAAACC CATAGGAAAA 450
GGCGGGCAGC CTTAGTGAGC TTCCTGCGGG GAGAGGTTTT TCTGTTAGAG 500
CTGGCANGGT CTGCTCATCG ACCATCTTCA GGCCTCGTGC C 541
```